U d'/ of Ottawa
39003001561645

Digitized by the Internet Archive
in 2011 with funding from
University of Toronto

http://www.archive.org/details/lalinguistiqu00hove

62209

BIBLIOTHÈQUE

DES

SCIENCES CONTEMPORAINES

BIBLIOTHÈQUE DES SCIENCES CONTEMPORAINES

LA LINGUISTIQUE

PAR

ABEL HOVELACQUE

LINGUISTIQUE. PHILOLOGIE. ÉTYMOLOGIE
LA FACULTÉ DU LANGAGE ARTICULÉ, SA LOCALISATION,
SON ORIGINE, SON IMPORTANCE DANS L'HISTOIRE NATURELLE.
CLASSIFICATION ET DESCRIPTION DES DIFFÉRENTS IDIOMES.
PLURALITÉ ORIGINELLE ET TRANSFORMATION
DES SYSTÈMES DE LANGUES.

QUATRIÈME ÉDITION

REVUE ET AUGMENTÉE

PARIS
LIBRAIRIE SCHLEICHER FRÈRES
8, RUE MONSIEUR-LE-PRINCE, 8

Tous droits réservés.

ANNEXE DE LA BIBLIOTHÈQUE
Universitas BIBLIOTHECA Ottaviensis
uOttawa LIBRARY ANNEX

P
121
.H68
1911

TABLE DES MATIÈRES

Pages

INTRODUCTION IX

AVERTISSEMENT POUR LA SECONDE ÉDITION..... XIII

AVERTISSEMENT POUR LA QUATRIÈME ÉDITION..... XV

CHAPITRE I. Linguistique, philologie étymologie..... 1
§ 1. Distinction de la linguistique et de la philologie 1
§ 2. La vie des langues..... 9
§ 3. Aide que se prêtent mutuellement la linguistique et la philologie..... 12
§ 4. Les polyglottes..... 14
§ 5. Les dangers de l'étymologie..... 16

CHAPITRE II. La faculté du langage articulé. Sa localisation. Son importance dans l'histoire naturelle 22

CHAPITRE III. Première forme linguistique. Le monosyllabisme. Les langues isolantes..... 39
§ 1. Le chinois..... 43
§ 2. L'annamite 52
§ 3. Le siamois 53
§ 4. Le birman..... 55
§ 5. Le tibétain..... 55

CHAPITRE IV. Seconde forme linguistique. L'agglutination. Les langues agglutinantes..... 61
§ 1. Qu'est-ce que l'agglutination ?..... 61
§ 2. Langues de l'Afrique méridionale.... 67
Langue des Hottentots..... 68
Langues des Bochimans..... 72
§ 3. Langues des nègres d'Afrique..... 73
§ 4. Langues du groupe bantou..... 83
§ 5. Le poul..... 91

Pages

CHAPITRE IV. § 6. Les langues nubiennes..... 93
§ 7. Langues des Négritos..... 94
§ 8. Langues des Papous..... 94
§ 9. Langues australiennes..... 95
§ 10. Langues maléo-polynésiennes..... 96
§ 11. Le japonais..... 105
§ 12. Le coréen..... 110
§ 13. Langues dravidiennes..... 111
§ 14. Langues ouralo-altaïques..... 125
I. Le groupe samoyède..... 127
II. Le groupe finnois..... 128
III. Le groupe turc..... 142
IV. Le groupe tongouse..... 148
V. Le groupe mongol..... 150
VI. De l'harmonie vocalique et de la parenté des langues ouralo-altaïques 152
§ 15. La langue basque..... 156
§ 16. Langues américaines 173
§ 17. Langues hyperboréennes..... 189
§ 18. Langues du Caucase..... 190
§ 19. De quelques autres idiomes classés parmi les langues agglutinantes.... 193
a. La prétendue langue scythique..... 193
b. La langue de la 2e colonne des inscriptions cunéiformes 194
c. La langue dite sumérienne ou accadienne 197
§ 20. La théorie des langues touraniennes. 201

CHAPITRE V. Troisième forme linguistique. La flexion... 205
§ 1. Qu'est-ce que la flexion ?..... 206
§ 2. Flexion indo-européenne et flexion sémitique 207
A. Les langues sémitiques..... 211
§ 1. Du sémitisme en général et de l'ensemble des langues sémitiques..... 212
§ 2. Groupe araméo-assyrien 218
I. Chaldéen et syriaque..... 218
II. Assyrien 221
§ 6. Le groupe chananéen..... 225
I. Hébreu 225
II. Phénicien 231
§ 4. Le groupe arabe..... 234
I. Arabe 234
II. Langues de l'Arabie méridionale et de l'Abyssinie..... 240

Pages

CHAPITRE V. § 5. Individualité des langues sémitiques. Leur patrie primitive 242

B. Les langues khamitiques 246

§ 1. Le groupe égyptien 248

§ 2. Le groupe libyen 253

§ 3. Le groupe éthiopien 255

C. Les langues indo-européennes 257

La langue commune indo-européenne. 260

§ 1. Branche hindoue 270

I. Anciennes langues hindoues ... 271

II. Langues néo-hindoues 276

III. Dialectes des Tsiganes 278

§ 2. Branche éranienne 279

I. Zend 280

II. Perse 284

III. Arménien 286

IV. Husvârèche 288

V. Parsi 291

VI. Persan 292

VII. Ossète, kourde, beloutche, afghan, etc 293

§ 3. Branche hellénique 295

§ 4. Branche italique 306

I. Anciennes langues italiques 307

II. Langues novo-latines 318

§ 5. Branche celtique 338

§ 6. Branche germanique 347

I. Gothique 349

II. Langues scandinaves 352

III. Bas allemand 354

IV. Haut allemand 360

§ 7. Branche slave 366

§ 8. Branche lettique 389

I. Lithuanien 390

II. Lette 392

III. Vieux prussien 393

§ 9. Langues indo-européennes non classées 394

I. Etrusque 395

II. Dace 397

III. Langues indo-européennes de l'Asie Mineure 398

IV. Langues indo-européennes dites scythiques 400

V. Albanais 400

Pages

CHAPITRE V. § 10. Du mode de subdivision de la langue commune indo-européenne et de la région où elle fut parlée..... 402

CHAPITRE VI. Pluralité originelle des langues et transformation des systèmes linguistiques..... 413
§ 1. Comment se reconnaît la parenté des langues 413
§ 2. Pluralité originelle des systèmes linguistiques 419
§ 3. Dans la vie historique les langues peuvent ne plus correspondre aux races. 422
§ 4. La transformation des espèces en linguistique 425

TABLE ANALYTIQUE 441

INTRODUCTION

Il était réservé aux dernières années du dix-huitième siècle de donner le jour aux méthodes d'investigation scientifique. La tâche était immense ; mais les hommes qui tentèrent de l'accomplir n'y firent pas défaut. C'est aux écrivains de l'*Encyclopédie* qu'il fut donné d'ouvrir l'ère contemporaine, la période de la science expérimentale.

L'esprit méthodique renouvela les procédés de recherche et les modes d'enseignement. Les sciences mathématiques, les sciences chimiques, les sciences naturelles rompirent enfin, et pour jamais, avec la métaphysique.

La *linguistique* n'est ni la moins importante ni la moins intéressante des sciences contemporaines ; ce volume lui est consacré. Notre intention est de montrer quelle place elle occupe dans l'histoire naturelle de l'homme. Tout d'abord, nous aurons à la définir. Les questions les plus

délicates de cette science sont abordées et résolues chaque jour par des personnes tout aussi ignorantes de son objet que de sa méthode. C'est le sort commun de toutes les sciences naturelles. On y supplée volontiers par des assertions purement sentimentales au défaut d'études fondées sur l'expérience. C'est ainsi qu'on se déclare hardiment polygéniste ou monogéniste, ami ou ennemi de la doctrine de l'évolution, sans avoir jamais mis le pied dans un laboratoire d'anthropologie.

Nous ne chercherons pas à éviter l'examen de la question de l'origine du langage. C'est une question purement anthropologique. Sans nous occuper des rêveries auxquelles elle a donné lieu, nous la traiterons uniquement au point de vue de l'histoire naturelle, c'est-à-dire de l'anatomie et de la physiologie. Le langage articulé est un fait naturel, soumis, comme tout autre fait, à l'investigation libre et désintéressée, et ce n'est pas une entreprise téméraire que d'aborder la question de son origine. L'écarter sous prétexte qu'il faut proscrire toute recherche des « origines premières », c'est admettre la possibilité même de ces causes premières, dont les mathématiques et la chimie ont fait justice.

A côté des questions de linguistique pure, nous

avons introduit çà et là, mais dans une faible mesures, certaines questions de philologie qui s'y rattachaient directement. Nous avons traité plus volontiers de quelques points d'ethnographie linguistique, mais d'une façon très-incomplète. Nous nous promettons d'y revenir. Quant aux questions de linguistique proprement dite, nous étions contraint, par la nature et le but de cette *Bibliothèque*, à les parcourir toutes fort rapidement ; c'est une difficulté dont le lecteur voudra bien tenir compte.

Avant d'entrer en matière, qu'il nous soit permis d'adresser nos remerciements à MM. Picot et Vinson, pour la part qu'ils ont prise à notre travail. Nous leur devons beaucoup : des notes, des renseignements, et surtout les conseils d'esprits sûrs et méthodiques.

AVERTISSEMENT

POUR LA SECONDE ÉDITION

C'est une édition véritablement nouvelle que nous publions aujourd'hui, et non pas une réimpression.

La partie des langues monosyllabiques et celle des langues à flexion n'ont reçu que des modifications peu importantes, mais il n'en a pas été de même de la partie des langues agglutinantes.

Nous avons développé tout particulièrement les chapitres relatifs à la langue des Hottentots, aux langues de la Guinée, et aux langues du système bantou.

Nous avons également revu et remanié le chapitre consacré aux idiomes de l'Amérique. On a fait récemment des progrès considérables dans la science de l'américanisme, mais tout est loin d'avoir été dit à ce sujet. Après avoir reproduit la classification assez communément acceptée des langues américaines, nous avons tâché de mettre en évidence ce fait très important que les langues

en question ne constituent pas une espèce particulière, et que leurs procédés se retrouvent dans un grand nombre d'autres langues agglutinantes.

En ce qui concerne d'autres idiomes également agglutinatifs, dont la phonétique, la structure et le vocabulaire demeurent encore fort obscurs, nous avons persisté dans notre première réserve. On peut ne pas approuver ; mais on blâmerait avec bien plus de droit des conclusions légères et trop hâtives.

Quant à la méthode qui nous avait guidé dans la rédaction de notre première édition, et qui est commune aux différents volumes de la *Bibliothèque des sciences contemporaines*, nous l'avons suivie avec toute la fidélité possible dans cette nouvelle publication. C'est la méthode expérimentale, la méthode sur laquelle reposent toutes les recherches et toutes les découvertes de l'esprit moderne.

AVERTISSEMENT

POUR LA QUATRIÈME ÉDITION

Nous avons à signaler, depuis l'époque à laquelle a paru la troisième édition de ce livre, la publication des travaux importants des « néo-grammairiens » sur la phonétique du système indo-européen ancien ; puis l'achèvement du grand ouvrage de Frédéric Müller, *Grundriss der Sprachwissenschaft*.

LA LINGUISTIQUE

CHAPITRE PREMIER.

LINGUISTIQUE — PHILOLOGIE — ÉTYMOLOGIE.

§ 1. Distinction de la linguistique et de la philologie.

Il est rare que dans le langage courant, et même dans les écrits scientifiques, on établisse une distinction entre les deux mots de *linguistique* et de *philologie ;* on les emploie d'ordinaire l'un pour l'autre, à peu près au hasard et selon qu'il faut satisfaire aux besoins euphoniques d'une phrase ou d'une période. Nos meilleurs écrivains, des érudits même, confondent sans cesse ces deux termes ; la philologie, la linguistique ne sont trop souvent, pour eux, que l'étude des étymologies, et ils donnent indifféremment aux personnes qui se livrent à cette sorte de recherches le nom de *linguistes* ou de *philologues.* L'examen de la parenté possible de deux idiomes australiens et la correction d'un texte de Plaute seraient indistinctement des travaux de linguistique ou de philologie.

Il est loin d'en être ainsi, et nous devons nous attacher, avant tout, à combattre cette grave erreur.

La linguistique est une science naturelle, la philologie une science historique.

Dans le Dictionnaire de la langue française d'Em. Littré, au mot *Linguistique*, nous lisons : « Etude des

langues considérées dans leurs principes, dans leurs rapports, et en tant qu'un produit involontaire de l'esprit humain ». Cette définition a un grand mérite : celui de ne pas s'appliquer tout aussi bien au mot *Philologie*. A ce dernier mot E. Littré donne trois sens divers : 1° Sorte de savoir général qui regarde les belles-lettres, les langues, la critique, etc... 2° Particulièrement : étude et connaissance d'une langue en tant qu'elle est l'instrument ou le moyen d'une littérature. 3° Philologie comparée ; étude appliquée à plusieurs langues, que l'on éclaire par la comparaison entre les unes et les autres. » De ces trois applications, les deux premières sont exactes, mais à propos de la dernière nous devons faire une réserve. L'auteur y définit d'une façon très heureuse la Philologie comparée ; mais le moyen de concevoir que la Linguistique puisse en aucun cas recevoir ce nom de Philologie comparée ? C'est avec juste raison que E. Littré distingue la Philologie simplement dite d'avec la Linguistique, mais il cède sans motif suffisant à l'usage qui fait dévier de son sens le terme de *Philologie*, alors qu'on lui applique l'épithète de *comparée*.

Comment, pour être comparée, la philologie se transformerait-elle en linguistique ? Nous avons peine à le comprendre. La physiologie comparée, celle, par exemple, qui embrasse les relations des végétaux et des animaux, n'aurait-elle plus droit au nom de *physiologie* ? L'anatomie comparée des diverses races humaines, ou, si l'on veut, l'anatomie comparée de l'homme et des autres primates, devrait-elle perdre le nom d'*anatomie* ?

Il en est évidemment de la philologie comme de ces autres sciences, et l'on ne saurait à aucun titre, lorsqu'elle devient comparée, ou, pour mieux dire, comparative, lui enlever son propre et véritable nom.

Rollin définissait les philologues « ceux qui ont tra-

vaillé sur les anciens auteurs pour les examiner, les corriger, les expliquer et les mettre au jour ». Cette définition conserve encore toute sa valeur ; elle correspond aux deux premiers sens que Littré comme nous venons de le voir, donne dans son Dictionnaire au mot de *Philologie*. En définitive, la tâche du philologue est l'étude critique des littératures sous le rapport de l'archéologie, de l'art, de la mythologie ; c'est la recherche de l'histoire des langues et subsidiairement de leur extension géographique ; c'est la découverte des emprunts qu'elles se sont faits les unes aux autres dans le cours des temps, en particulier des emprunts lexiques ; c'est, enfin, la restitution et la correction des textes.

C'est là, au premier chef, une science historique, une branche considérable de l' « érudition ». Avant le développement contemporain des sciences naturelles, les langues n'étaient envisagées, et il n'en pouvait être autrement, que sous ce seul et unique rapport ; la philologie a précédé de longtemps la linguistique.

La philologie, simplement dite, ne s'attache qu'à une seule langue ; elle la critique, en interprète les documents, en améliore les textes d'après les données et les informations que peut lui fournir cette seule et même langue. L'étude vient-elle à se porter de façon corrélative sur deux langues diverses, ou sur plusieurs branches d'un même idiome, la philologie devient alors comparée. Ainsi la philologie dite *classique* est le plus souvent comparée : elle s'occupe, comme l'on sait, des textes grecs et latins. De même la philologie romane, la philologie germanique, la philologie slave, sont, les unes et les autres, comparées ; elles traiteront, par exemple, de l'influence qu'exerça la langue des Précieuses du dix-septième siècle sur la langue courante des âges suivants ; du rôle que joua dans la formation de l'allemand moderne la version de la Bible par Luther ; de

l'extension des langues slaves, vers l'ouest de l'Europe, au moyen âge, puis de leur rétrogradation vers l'est. Egalement *comparée* est la philologie dite *orientale* qui s'applique à ces trois langues, le persan, l'arabe, le turc, tout étrangères que soient les unes aux autres ces différentes langues sous le rapport linguistique. Dans l'Inde et dans l'extrême Orient le bouddhisme a donné naissance à une philologie comparée, tout comme la légende de Charlemagne dans l'Europe occidentale.

C'est en particulier à Schleicher (1), c'est à Kuhn, Chavée (2), Spiegel (3) qu'est due la distinction si importante entre ces deux sciences, philologie et linguistique. Tous ces auteurs tombent d'accord sur le fait capital que l'une est du domaine des connaissances historiques, l'autre du domaine des connaissances naturelles.

La linguistique peut être définie ; l'étude des éléments constitutifs du langage articulé et des formes diverses qu'affectent ou peuvent affecter ces éléments. En d'autres termes, si l'on veut, la linguistique est la double étude de la phonétique et de la structure des langues.

Il est aisé de comprendre comment la linguistique se rattache à la physiologie par l'étude du matériel phonétique des langues, c'est-à-dire de leurs sens. Le premier soin du linguiste est d'inventorier les voyelles et les consonnes des langues qu'il examine et d'établir les lois de leurs permutations ou de leurs variations ; la découverte de ces lois lui sera d'autant plus facile qu'il sera plus familiarisé avec le jeu de l'appareil vocal.

Les voyelles et les consonnes constituent les premiers éléments du langage. Plus tard apparaissent d'autres éléments, que l'on qualifie souvent du nom d'éléments simples bien que, pour l'ordinaire, ils soient déjà com-

(1) *Die deutsche sprache*, Intr., chap. VI.
(2) *Bulletins de la Société d'anthropologie de Paris*, 1862, p. 198.
(3) *Die traditionnelle literatur der Parsen*, p. 48.

posés (c'est-à-dire formés d'un ensemble de voyelles et de consonnes) ; ce sont les monosyllabes auxquels on donne le nom de *racines*.

Ces monosyllabes, l'expérience nous les fera découvrir au fond de tous les systèmes linguistiques. Tantôt ils seront formés d'un seul élément sonore, c'est-à-dire d'une seule voyelle ; tantôt ils seront formés de la réunion de plusieurs éléments : soit, par exemple une consonne suivie d'une voyelle, soit une voyelle suivie d'une consonne, soit encore quelque autre groupement. Mais la signification de ces premières racines ne sera jamais que très générale et elles se trouveront étrangères à toute notion de genre, de cas, de nombre, de personnes, de temps, de modalité.

L'étude de ces éléments constitue, disons-nous, l'un des premiers soins du linguiste. En second lieu arrive l'examen des formes qu'affectent ou peuvent affecter ces éléments ; cette nouvelle étude reçoit le nom de *morphologie*. Nous traiterons plus loin des différentes variétés morphologiques du langage, c'est-à-dire des différents modes de structure que peuvent présenter les langues, et nous constaterons alors que des idiomes qu'il convient de ranger, sous ce rapport, dans un seul et même groupe, par exemple les langues dites *agglutinantes*, peuvent être, si l'on envisage leurs éléments constitutifs, étrangères de tous points les unes des autres. C'est ainsi que les langues indo-européennes et les langues sémitiques, dont les racines sont tout à fait différentes, tout à fait irréductibles, se trouvent les unes et les autres dans la même classe morphologique ; de même, le turc, le basque, le japonais, le tamoul ont, en général, la même structure ; mais les radicaux de ces différentes langues sont essentiellement différents, et il est impossible de les ramener scientifiquement à une origine commune, à une seule et même souche.

Ce sujet nous occupera en son temps comme il le mérite. Notre but, pour l'instant, est de bien établir ce fait capital, que la linguistique appartient au groupe des sciences naturelles, et que pour la ranger parmi les sciences historiques il faut méconnaître à la fois et son but et sa méthode.

C'est à Auguste Schleicher que nous devons les écrits les plus nets et les plus démonstratifs sur cet important sujet.

Schleicher, chose rare parmi ses compatriotes, était un esprit parfaitement dégagé d'aspirations métaphysiques. Il avait traversé, comme tant d'autres, les écoles transcendantes ; comme tant d'autres, il avait suivi les docteurs du théurgisme et de l'hyperphysisme, mais leurs fantaisies n'avaient pu séduire cette intelligence positive et peu disposée à se payer d'emphatiques et vaines paroles. Schleicher était l'homme de l'expérience, l'homme de la méthode. C'est à lui le premier, comme l'on sait, qu'il échut de dresser un code général de la phonétique et de la structure des langues indo-européennes. W. Jones, vers la fin du XVIIIe siècle, avait définitivement affirmé la parenté de ces langues, et Bopp, au commencement du XIXe, avait démontré méthodiquement cette même parenté.

Ainsi qu'il aimait à le dire lui-même, il est certain que ses remarquables connaissances en botanique lui furent d'une utilité capitale pour ses recherches sur la morphologie des langues, tant les procédés d'analyse et de comparaison sont identiques dans l'étude de toutes les sciences naturelles.

L'ingénieuse analogie que, pour bien faire comprendre la distinction de la linguistique d'avec la philologie, Schleicher se plut à établir entre le linguiste et le botaniste d'une part, et d'autre part le philologue et l'horticulteur, mérite à tous égards d'être rappelée. On

la trouve dans son excellent livre sur la langue allemande (1) :

« La philologie, disait-il, est une science historique, et cette science ne peut se trouver appliquée que là où l'on est en présence d'une littérature, d'une histoire. Là où les monuments font défaut, là où il n'y a point de culture littéraire, le philologue n'a que faire ; la philologie, en un mot, ne peut s'exercer que sur des documents historiques. Il en est tout différemment de la linguistique, dont l'objet unique est la langue elle-même, dont l'unique étude est l'examen de la langue en elle-même et pour elle-même. Les variations historiques des langues, le développement plus ou moins factice de leur vocabulaire, souvent même leurs procédés syntactiques, tout cela n'est pour le linguiste que d'une importance secondaire ; il consacre son soin tout entier à l'étude de la manifestation elle-même du langage articulé, fonction naturelle, inévitable et déterminée, à laquelle l'homme ne pourrait se soustraire, et qui, ainsi que toutes les autres fonctions, est d'une implacable nécessité. Peu importe au linguiste qu'une langue ait régné, des siècles durant, sur de vastes empires, qu'elle ait donné naissance aux monuments littéraires les plus glorieux, qu'elle se soit prêtée aux exigences de la culture intellectuelle la plus délicate, la plus raffinée ; peu lui importe aussi qu'une langue obscure ait misérablement péri, sans fruits, sans rejetons, étouffée par d'autres idiomes, inconnue à jamais du philologue. La littérature est, sans conteste, un auxiliaire puissant grâce auquel il est aisé de saisir l'idiome lui-même, de reconnaître la succession de ses formes, les phases de son développement ; un auxiliaire précieux, mais non pas indispensable. Ajoutez que la connaissance d'une seule langue ne peut suffire au linguiste, et en cela il se distingue encore du philo-

(1) *Die deutsche sprache*, Introduction.

logue. Il existe, par exemple, une philologie latine, tout indépendante de la philologie grecque : une philologie hébraïque, tout indépendante de la philologie arabe ou assyrienne. Mais il ne saurait être question d'une linguistique purement latine, d'une linguistique purement hébraïque : la linguistique est comparée ou n'est pas. On ne peut en effet se rendre compte d'une forme qu'en la comparant à d'autres formes. La philologie peut donc être spéciale, particulière à un seul idiome ; mais lorsqu'il s'agit d'étudier les éléments constitutifs d'une langue et sa structure, il faut déjà connaître la phonétique et la structure d'un certain nombre d'autres idiomes. Répétons-le donc une fois encore, les recherches du linguiste sont toujours et essentiellement comparatives, à l'encontre de celles du philologue, qui peuvent être toutes spéciales. »

C'est ici que Schleicher place son ingénieuse et très-juste comparaison : « Le linguiste, dit-il, est un naturaliste ; il étudie les langues à la façon dont le botaniste étudie les plantes. Le botaniste doit embrasser d'un coup d'œil l'ensemble des organismes végétaux ; il recherche les lois de leur structure, celles de leur développement, mais il ne se préoccupe en aucune manière du plus ou moins de valeur des plantes, de leur usage plus ou moins précieux, de leur agrément plus ou moins reconnu. A ses yeux, la première venue des mauvaises herbes peut avoir un bien autre prix que n'en ont les roses les plus belles, les lis les plus rares. Le rôle du philologue est tout différent. Ce n'est point au botaniste, mais bien à l'horticulteur qu'il convient de le comparer. Ce dernier ne donne ses soins qu'à telles ou telles espèces, qui sont l'objet d'une faveur particulière ; c'est la beauté de la forme qu'il recherche, c'est la coloration, c'est le parfum. Une plante inutile est sans valeur à ses yeux ; il n'a que faire des lois de la structure et du développement : le végétal

qui, sous ce rapport, put posséder la valeur la plus considérable a chance de n'être pour lui qu'une mauvaise herbe vulgaire. »

Cette comparaison est exacte, et, mieux que toute autre explication, elle dit assez que le linguiste étudie chez l'homme le phénomène du langage articulé et ses produits à la façon dont tout physiologiste étudie les autres fonctions, la locomotion, par exemple, l'olfaction, la vision, ou encore la digestion, la circulation. Et non-seulement il recherche et détermine les lois normales propres à ce phénomène, mais encore il découvre et caractérise les altérations véritablement pathologiques qui se présentent maintes fois durant le cours de la vie des langues.

§ 2. La vie des langues.

Les langues en effet naissent, croissent, dépérissent et meurent comme tous les êtres vivants. Elles ont passé tout d'abord par une période embryonnaire, elles atteignent un complet développement et sont livrées, en fin de compte, à la métamorphose régressive. C'est précisément cette conception de la *vie* des langues qui, ainsi qu'on l'a déjà remarqué, distingue la science moderne du langage d'avec les spéculations du passé.

Nous traiterons dans un autre chapitre de la naissance des langues et de l'origine de la faculté du langage articulé ; plus loin aussi nous verrons comment les systèmes linguistiques les plus compliqués proviennent de systèmes rudimentaires ; comment, en un mot, les formes dont l'organisation est la plus complète proviennent de formes beaucoup moins développées.

Les langues une fois nées, l'on ne peut dire qu'elles entrent aussitôt dans leur période historique, en entendant par là que leur développement se trouve soumis

d'ores et déjà à l'arbitraire et aux fantaisies de ceux qui les parlent. Ce serait là une erreur. Le développement des langues est, avant tout, déterminé, et le cours de leur vie ne saurait, par une inadmissible dérogation aux lois naturelles, échapper aux nécessités communes à tout ce qui vit. A la vérité, sous l'influence de circonstances heureuses ou malheureuses, elles s'altéreront plus ou moins gravement, elles marcheront à leur décadence, à leur perte, d'un pas plus ou moins précipité ; mais rien ne fera fléchir leurs tendances organiques.

Elles sont, en un mot, ce que leur nature veut qu'elles soient. Jamais, par exemple, on ne parviendrait à créer une langue mixte. On ne saurait imaginer une langue indo-européenne dont la grammaire soit en partie slave, en partie latine. Il n'y a pas, il ne peut y avoir de langues mixtes. L'anglais, par exemple, chez lequel se sont introduits un si grand nombre d'éléments étrangers, notamment d'éléments français, n'en demeure et n'en demeurera pas moins jusqu'à son extinction une vraie langue germanique ; le basque est dans un cas analogue : ses emprunts constants à deux langues romanes n'altéreront jamais son caractère particulier. C'est encore ainsi qu'au moyen âge le huzvârèche conserva son caractère de langue éranienne, en dépit de l'intrusion considérable d'éléments sémitiques dont il eut à souffrir.

Mais il ne faut point douter que cette sorte de commerce intellectuel, que ces emprunts, fruits inévitables de la civilisation, ne précipitent singulièrement la vie des langues. Les faits sont là, évidents, palpables. Ainsi, parmi les langues germaniques, nous voyons l'anglais parcourir du milieu du treizième siècle jusqu'à nos jours une rapide, très rapide carrière, tandis que l'islandais nous offre aujourd'hui encore et à chaque instant des formes très anciennes et fort bien conservées. L'obscur lithuanien peut être tenu pour le moins dégradé des

idiomes indo-européens de l'Europe, et, selon toute vraisemblance, il nous aurait fait admirer longtemps encore ses formes antiques et précieuses, si la rude concurrence de l'allemand ne le menaçait d'une disparition prochaine. C'est ainsi que périssent chaque jour dans des luttes inégales, mais que rien ne saurait prévenir, des êtres pleins de vie et de santé qui, à la faveur de circonstances moins funestes, auraient connu de longues années et ne se seraient pas éteints, misérables, sans postérité.

Il est difficile de supposer qu'un système linguistique arrivé à l'âge le plus florissant, le plus riche de son développement, n'entre pas aussitôt dans la période de métamorphose régressive, et il est tout aussi difficile que cette période ne soit point caractérisée d'une façon spéciale par la tendance de plus en plus individualiste des idiomes de ce système. Nous savons, par exemple, que les langues dites *indo-européennes* ou *aryennes* (hindoues, éraniennes, helléniques, italiques, celtiques, germaniques, slaves, lettiques) proviennent d'une mère commune, dont il a été possible de déterminer les éléments phonétiques et de restituer, au moins en ses traits essentiels, la morphologie, la structure : or, il est supposable que la période de formation prit fin au moment même où commencèrent à se manifester des divergences dialectales et qu'il n'y eut point d'intervalle sensible entre ces premiers temps et la période de métamorphose régressive. Le linguiste doit avant tout déterminer, ou pour mieux dire, restituer les formes qu'affectaient, au moment de leur division en dialectes, les langues mères dont il n'existe pas de monuments écrits. Ainsi que nous l'avons dit, la tâche se trouve assez avancée pour le système indo-européen ; mais elle est à peine ébauchée en ce qui concerne les langues sémitiques (chaldéen, syriaque, hébreu, phénicien, arabe, etc.) et est tout en-

tière à entreprendre pour le plus grand nombre des autres systèmes ; celui, par exemple, des langues dites khamitiques (ancien égyptien, copte, tamachek, galla, etc.) et celui des langues dravidiennes (tamoul, télinga, etc.).

Mais la vie des langues n'est point un sujet qu'il soit possible de traiter en quelques pages, il réclamerait un volume entier et une longue série d'exemples pris tour à tour dans les différentes familles linguistiques. Nous n'entamerons pas cet exposé trop spécial, et il suffira sans doute d'avoir signalé ici ce fait général. Cf. Whitney, *la Vie du langage*, Sayce, *Introd. to the Science of Language*. Voir également *Bulletins de la Soc. d'anthropologie*, 1885, p. 371.

§ 3. Aide que se prêtent mutuellement la linguistique et la philologie.

Il est incontestable que le linguiste trouve parfois un puissant auxiliaire dans l'emploi de la méthode historique. Cette dernière est indispensable en effet lorsqu'il s'agit de l'étude de la syntaxe. Ici l'initiative personnelle peut être plus marquée. Loin de nous, certes, la moindre velléité d'attribuer à cette initiative une liberté à laquelle elle ne saurait prétendre sans braver les premiers enseignements de l'expérience ; nous savons assez que la spontanéité est déterminée de la manière la plus stricte et que le prétendu libre arbitre n'est, selon la parole de Spinosa, que la conscience de la volonté. Il nous faut donc encore considérer cette sorte d'arbitraire comme le fruit, le simple fruit d'une disposition naturelle, soumise, par conséquent, à une direction également naturelle. L'on peut dire que les formations par analogie, elles-mêmes, n'échappent pas à ce sort commun et qu'elles ne trahissent, le plus souvent, qu'une véritable paresse intellectuelle.

Nous nous trouvons amené à répéter ici que la science naturelle de la linguistique et la science historique de la philologie ne sont point rivales l'une de l'autre et que rien ne saurait autoriser à les tenir pour deux sciences hostiles. En effet, deux ordres de connaissance, si distincts qu'ils soient, ne peuvent conduire à des résultats opposés, et deux véritables sciences, deux sciences vraiment dignes de ce nom, ne sauraient, en aucun cas, être ennemies l'une de l'autre. Les sciences au contraire se complètent mutuellement, et chacune d'elles est vis-à-vis des autres débitrice et créancière tout à la fois.

Tel est, en particulier, le cas de la linguistique et de la philologie. Le philologue doit connaître, au moins d'une façon générale, les résultats acquis par le linguiste. S'il ne sait rien de la langue elle-même, de cet agent le plus considérable de la pensée, s'il ignore et sa structure et les éléments qui la composent, comment pourra-t-il porter quelque jugement complet sur les produits, sur les fruits de cet agent ? Autant dire qu'un ethnographe pourrait faire bon marché d'un ensemble de données élémentaires relatives à l'anatomie des races, et n'en tenir même aucun compte. C'est là une considération presque banale, et pourtant il est bon nombre de philologues qu'elle n'a point le don de satisfaire. De là cet amas de dissertations subjectives, sans but, sans doctrine, ce fatras d'arguties oiseuses où la rhétorique le dispute au vide et à l'ineptie. Les librairies françaises, par un reste de chance heureuse, n'en sont pas les plus encombrées.

Le philologue, par contre, prépare au linguiste un matériel précieux. Il lui facilite la connaissance des formes historiques du langage et lui expose ce qu'il a pu découvrir de leur chronologie et de leur succession ; il lui découvre enfin les divergences dialectales d'où peuvent sortir tant et de si précieuses instructions.

Si donc il importe de distinguer ces deux sciences, de ne confondre ni leur but, ni leur méthode, pas plus que leur vrai nom, il n'importe pas moins de reconnaître qu'elles sont appelées l'une et l'autre à se rendre des services mutuels et considérables. C'est ainsi que l'histoire a maintes fois fourni à l'étude des races humaines d'utiles informations et que l'anthropologie, à son tour, a pu éclaircir bien des faits historiques.

§ 4. Les polyglottes.

La connaissance pratique des langues, ou, pour nous exprimer d'une manière plus simple, l'art de les parler couramment et de façon correcte, repose avant tout sur une aptitude naturelle. Cette aptitude se développe par un usage plus ou moins prolongé ; mais il ne serait exact, en aucun cas, de la regarder comme une science. L'on s'étonne souvent de voir un auteur de nombreux et bons travaux linguistiques être peu capable d'entretenir la conversation en quatre ou cinq langues différentes, et l'on est tout surpris qu'il ne sache se servir parfois, avec quelque facilité, que de son idiome maternel. Il y a là une forte méprise. Le linguiste n'a que faire d'être polyglotte, ou, du moins, il n'est point nécessaire qu'il le soit. Le polyglotte, de son côté, n'a, du fait même de son art, aucun droit au nom de linguiste ; et cependant chaque jour nous entendons donner ce nom de linguistes aux personnes qui, grâce à certaines circonstances, grâce notamment à cette aptitude spéciale, parlent avec plus ou moins de facilité dix, douze idiomes, parfois même davantage, sans connaître cependant un traître mot de leur structure. Ce que nous avons dit plus haut du caractère même de la linguistique et de la nature des études du linguiste nous dispense d'insister sur cette confusion vulgaire.

Nous pensons toutefois que les résultats de la linguistique peuvent faciliter, jusqu'à un certain point, l'étude de l'art dont il s'agit. Prenons, par exemple, les langues romanes, issues, comme l'on sait, du latin vulgaire ; il est incontestable que l'on peut passer de l'une à l'autre d'après des règles à peu près fixes, en ce qui concerne particulièrement la phonétique, surtout en ce qui a trait à l'équivalence des consonnes. Un très petit nombre de principes généraux donnent la clef des concordances les plus communes ; la ressemblance des mots italiens, espagnols, français n'est plus fortuite ; elle devient, au contraire, logique, rationnelle, et leur étude marche d'un pas d'autant plus rapide qu'elle est moins abandonnée au hasard.

Les langues germaniques, elles aussi, possèdent des lois d'équivalence tout aussi précises ; à telles ou telles consonnes de l'allemand, par exemple, répondent telles ou telles consonnes de l'anglais, du hollandais, du suédois. Il en est de même pour les langues slaves : le tchèque, le russe, le croate ont une phonologie parfaitement fixe qui permet de passer sans peine des formes de l'un de ces idiomes aux formes de ses congénères. Répétons-le, il n'est pas besoin d'efforts intellectuels considérables pour atteindre à ce résultat ; il suffit de la connaissance de quelques principes élémentaires.

Nous ne nous illusionnons pas sur le peu de succès que l'on pourrait obtenir en introduisant dans l'instruction secondaire quelques notions de grammaire comparée. Il est difficile qu'un élève de dix, douze ou quinze ans s'intéresse d'une façon suivie aux lois de la permutation des consonnes et des voyelles dans les langues qu'il étudie ; il cherche à apprendre le grec et le latin comme il a appris sa langue maternelle, par la pratique pure et simple et sans s'occuper des règles formulées plus ou moins savamment. Mais n'y aurait-il pas un

grand bénéfice à ce que ceux-là au moins qui ont la charge de l'enseignement sussent que ces règles existent et n'ignorassent point les principales ni les plus élémentaires d'entre elles ? A notre sens, ce ne serait pas trop demander.

§ 5. Les dangers de l'étymologie.

Si l'aptitude spéciale à la connaissance pratique des langues n'est point une science, l'*étymologie*, par contre, telle qu'elle est pratiquée le plus souvent, ne peut être regardée ni comme une science ni comme un art. L'étymologie, par elle-même, n'est qu'une jonglerie, une sorte de jeu d'esprit, si bien que le grand ennemi de l'étymologiste, son ennemi implacable, c'est le linguiste. En un mot, l'étymologie par elle-même et pour elle-même n'est que de la divination ; elle fait abstraction de toute expérience, néglige les difficultés et se contente des apparences spécieuses de ce qui n'est qu'à peine probable ou à peine vraisemblable. Peut-on, à première vue, et de prime abord, douter que les mots de l'allemand moderne *bereit* « prêt », *œhnlich* « analogue, semblable », *abenteuer* « aventure » ne répondent au latin *paratus*, au grec ἀνάλογος, au français *aventure ?* L'anglais *to call* au grec καλέω « j'appelle, je convoque ». Et cependant il n'en est rien.

L'analyse linguistique démontre l'inanité de ces rapprochements faciles ; ils ne soutiennent pas une seconde l'examen d'une critique méthodique. C'est à l'aide de procédés aussi fantaisistes que l'on a prétendu assimiler les idiomes absolument étrangers les uns aux autres, les langues sémitiques et les langues indo-européennes, le basque et l'irlandais. Les plus illustres sémitisants, ceux qui ont rendu à la philologie des langues syro-arabes les meilleurs services, se sont maintes fois laissé prendre à ce piège, et nous voyons à tout instant

dans leurs écrits des racines sémitiques et des racines indo-européennes rapprochées sans critique les unes des autres. Gesenius lui-même n'a point échappé à ce malentendu, et il n'est pas étonnant qu'à sa suite les éxégètes orthodoxes y aient donné à cœur joie. Rien de plus périlleux que de s'emparer de deux mots tout faits et de les rapprocher l'un de l'autre, si l'on ignore les procédés et les lois de leur structure ; les équivalences qui semblent au premier coup d'œil s'imposer le plus invinciblement sont parfois les plus trompeuses. Bien souvent, au contraire, des formes que l'on ne songeait jamais à rapprocher les unes des autres se trouvent unies par les liens de la plus étroite parenté. Depuis leur antique communauté, depuis l'époque où elles n'étaient toutes qu'une seule et même forme, elles ont subi chacune des lois diverses de variation ; mais ces lois sont découvertes aujourd'hui, et l'unité, la réelle unité de ces formes, est un fait hors de conteste. C'est ainsi, par exemple, que le grec ἡδύς « doux » et le latin *suavis* remontent tous deux à une seule et même forme plus ancienne ; il en est de même du latin *solus* et du perse *haruva* « tout », de l'irlandais *il* et du sanskrit *purus* « nombreux » ; du grec ἰός « poison » et du latin *virus*, de l'anglais *five* « cinq » et du croate *pet ;* du hollandais *vader* « père » et de l'arménien *hayr ;* de l'arménien *és* « je » et du croate *ja*. C'est ainsi encore que des mots appartenant à une seule et même langue et qui semblent, au premier abord, n'avoir aucune connexité, appartiennent en réalité à une seule et même racine ; en français, par exemple : solide, solder, soldat, seul, serf ; — jeu, bon, jour, divin ; — auspice, sceptique, évêque, épice, répit ; — assister, coûter, étable, obstacle. Nous sortirions des limites permises à cet écrit en exposant par le menu les principes qui relient entre elles ces formes diverses et pourtant proches alliées, que la pure et simple divination aurait

grand'peine sans doute à rattacher les unes aux autres.

Qu'est-ce donc que l'étymologie, ou plutôt que doit-elle être pour mériter créance et prétendre à une valeur scientifique ? Un résultat pur et simple. Résultat de la linguistique, résultat de la philologie.

Elle est déductive dans le premier cas, historique dans le second.

Disons quelques mots de ces deux hypothèses, en commençant par la seconde. L'histoire de la langue française nous enseigne, pour prendre quelques exemples, que *dinde* est un abrégé de *poule d'Inde ;* que les mots *guise*, *fauteuil*, *meurtre*, *heaume* sont d'origine germanique ; que les mots *alouette*, *cervoise*, *arpent* sont d'origine celtique. Voilà tout autant d'exemples d'étymologies philologiques, ou, si l'on veut, historiques. Sur ce terrain, en effet, c'est à la critique historique, à elle seule, qu'il appartient de décider si les suppositions que l'on se plaît à faire sont exactes ou inexactes, si elles sont vraisemblables ou invraisemblables. Mais la critique historique a trop souvent été en défaut. C'est de la critique historique que relèvent une foule d'étymologies appuyées sur des *parce que*, et dans le nombre il s'en rencontre plus d'une qui, pour paraître très simple au premier coup d'œil, n'en doit pas moins être regardée comme absolument défectueuse. Ainsi, d'après les juristes latins, l'esclave, *servus*, tirait son nom de ce qu'il avait été, par la grâce du vainqueur, sauvé, préservé d'un coup fatal ; or, tout au contraire, le sens antique de ce mot est celui de *protecteur*, de *gardien :* il répond rigoureusement, en tant que nominatif singulier, à la forme *haurvô*, gardien (*paçus-haurvô*, gardien de bétail) de l'Avesta. C'est à l'aide de *parce que* que l'on fait venir *feu* (défunt) de *fuit*, il fut. Un pas de plus et l'on tire *cadaver* de *ca* [ro] *da* [ta] *ver* [mibus], *nobilis* de *non vilis* et *dignus* de *di-genus*, espèce de dieu.

L'étymologie linguistique est tout aussi périlleuse, plus périlleuse peut-être, que l'étymologie philologique. « Sais-tu bien, demande le docteur, d'où vient le mot de *galant homme ?* — LE BARBOUILLÉ. Qu'il vienne de Villejuif ou d'Aubervilliers, je ne m'en soucie guère. — LE DOCTEUR. Sache que le mot de *galant homme* vient d'*élégant ;* prenant le *g* et l'*a* de la dernière syllabe, cela fait *ga*, et puis prenant *l*, ajoutant un *a* et les deux dernières lettres, cela fait *galant*, et puis ajoutant *homme*, cela fait *galant homme.* » Les moins mauvaises des étymologies de cette sorte — si tant est que toutes ne se vaillent point — sont peu supérieures à celles-là, soit dit sans exagérer. Il n'est pas plus rationnel, par exemple, de rapprocher le grec μορφή « forme, figure, aspect » et le latin *forma*, en prétendant que les consonnes *m* et *f* ont simplement changé de place, qu'il ne l'est de tirer *galant homme* d'*élégant.* La consonne *f* du latin, placée au commencement des mots, répond, comme nous le verrons plus loin, à une explosive aspirée (bh, dh ou gh) de la forme indo-européenne commune ; dans le cas actuel c'est à un « dh » que reproduit précisément le mot sanskrit *dharma-* dont le sens est celui de « jus, justitia ». On connaît le diminutif latin du mot *forma* qui est *formula* « forme, formule, précepte ». Quant à μορφή il est apparenté à μάρπτω « je saisis ».

Combien de personnes trouvent parfaitement vraisemblable cette prétendue et fausse équivalence du latin *forma* et du grec μορφή qui sont les premières à rire de Ménage, lorsqu'il tire *rat* du latin *mus* par l'entremise des formes soi-disant intermédiaires *muratus*, puis *ratus ?* Les deux étymologies pourtant se valent l'une l'autre.

C'est une idée trop répandue que celle de considérer le linguiste comme un faiseur d'étymologies, et ceux-là peuvent seuls entretenir cette illusion qui ne soupçonnent

ni le but ni la méthode de la linguistique. Aux yeux du linguiste, en effet, ces ressemblances plus ou moins fortes ne sont rien moins que déterminantes. L'expérience lui a fait connaître à quel point elles peuvent être trompeuses ; mais surtout, et avant tout, elle lui a appris que les langues ne sont pas des créations de hasard et qu'elles répondent, comme toute fonction, à une nécessité organique ; que les lois qui les régissent révèlent une précision d'autant plus éclatante qu'on les recherche avec plus de méthode ; que ces lois enfin découvrent et expliquent en maintes circonstances la parenté directe ou indirecte des mots, mais que la recherche de cette parenté n'est qu'un fait accessoire, un fait accidentel.

L'étymologiste, a-on dit, fait peu de cas des consonnes et néglige toutes les voyelles. Cela est parfaitement exact. L'étymologiste qui se livre à l'étymologie par elle-même et pour elle-même ignore de tout point ce que c'est que la philologie, et plus encore, s'il est possible, ce que c'est que la linguistique. Qu'un linguiste, qu'un philologue s'occupent d'étymologies, fort bien ; mais le privilège de cette sorte de recherches ne doit appartenir qu'à eux seuls. C'est avec les procédés de l'étymologie courante que l'on a fait du basque un parent de l'irlandais, du français ou du provençal un idiome celtique, du latin un dérivé du grec, du phénicien tout ce que l'on a voulu ; c'est avec l'étymologie pure et simple qu'aujourd'hui encore l'on prétend, à l'aide de quelques noms géographiques pris à peu près au hasard, caractériser la langue des anciens Ibères ; c'est avec cette même étymologie que l'on a lu couramment, en deux ou trois langues différentes, les inscriptions étrusques, que l'on pourrait encore les lire en une douzaine d'autres langues.

Nous ne saurions trop le répéter, la linguistique n'a rien de commun, ni de près ni de loin, avec ces exercices divinatoires. Le premier écueil dont elle garde ses disci-

ples, c'est la tentation de rapprocher des mots qui n'ont pas été au préalable méthodiquement analysés. A chaque instant l'étymologiste cède à cette tentation. Il n'opère, précisément, qu'au moyen de ces comparaisons aventureuses. Sans doute, le linguiste devra parfois se laisser guider par de pures et simples présomptions ; mais celles-ci ne pèseront ni sur ses conclusions ni sur le mode de ses recherches. Ce qu'il prétend découvrir, ce qu'il étudie, ce sont les éléments simples des langues et les procédés d'agrégation de ces éléments ; c'est le système de fonctionnement des formes organiques ; ce sont les lois qui président au développement de ces formes et ensuite à leurs altérations.

La linguistique n'est donc qu'une science naturelle.

C'est, d'ailleurs, ce que nous allons constater à nouveau en entrant dans un autre ordres d'idées.

CHAPITRE II.

LA FACULTÉ DU LANGAGE ARTICULÉ
SA LOCALISATION
SON IMPORTANCE DANS L'HISTOIRE NATURELLE

L'homme n'est homme que parce qu'il possède la faculté du langage articulé. C'était là jadis une proposition malsonnante. Elle est passée aujourd'hui à l'état de vérité banale, aux yeux du moins des personnes qui tiennent pour liquidé, et bien liquidé le compte de la métaphysique.

Sans doute, c'est un raisonnement peu convaincant que d'en appeler aux autorités, mêmes les plus reconnues ; pourtant il ne nous sera pas interdit de citer, à propos du sujet qui nous occupe, l'opinion de quelques auteurs dont la science s'honore à bon droit ; celle, par exemple, de Charles Martins : « Le langage articulé est le caractère distinctif de l'homme (1) » ; celle de Darwin : « Le langage articulé est spécial à l'homme, bien que, comme les autres animaux, il puisse exprimer ses intentions par des cris inarticulés, par des gestes et par les mouvements des muscles de son visage (2) » ; celle de Hunfalvy : « L'origine de l'homme doit être placée à l'origine du langage (3) ; celle de Hæckel : « Rien n'a dû ennoblir et transformer les facultés et le cerveau de l'homme autant que l'acquisition du langage. La diffé-

(1) *La création du monde organisé*, Revue des Deux-Mondes, 15 décembre 1871, p. 778.
(2) *La descendance de l'homme et la sélection sexuelle*, trad. franc. de E. Barbier, t. I, p. 53.
(3) Congrès international d'anthropologie et d'archéologie préhistoriques ; cinquième session, p. 436.

renciation plus complète du cerveau, son perfectionnement et celui de ses plus nobles fonctions, c'est-à-dire des facultés intellectuelles, marchèrent de pair, et en s'influençant réciproquement, avec leur manifestation parlée. C'est donc à bon droit que les représentants les plus distingués de la philologie comparée (c'est la *linguistique* que l'auteur a voulu dire) considèrent le langage humain comme le pas le plus décisif qu'ait fait l'homme pour se séparer de ses ancêtres animaux. C'est un point que Schleicher a mis en relief dans son travail *Sur l'importance du langage dans l'histoire naturelle de l'homme.* Là se trouve le trait d'union de la zoologie et de la philologie comparée ; la doctrine de l'évolution met chacune de ces sciences en état de suivre pas à pas l'origine du langage ». Et plus loin : « Il n'y avait point encore chez cet homme-singe de vrai langage, de langue articulée exprimant des idées (1). »

En temps et lieu nous reviendrons sur la corrélation de la naissance de l'homme et de celle de la faculté du langage articulé. Nous nous en tenons pour l'instant à ce point capital, que la faculté dont il s'agit constitue la caractéristique unique de *l'humanité.*

C'est en vain que l'on a cherché dans la comparaison de la constitution anatomique de l'homme et de celle des animaux inférieurs une divergence quelconque, un autre écart que celui du plus ou moins. Et cet écart a-t-il encore été diminué d'une façon considérable, à tous les yeux désintéressés, depuis la découverte des anthropoïdes africains. On peut dire que la théorie sentimentale du *règne humain* se trouve définitivement à bas et que son discrédit est parachevé. Ni l'évolution dentaire, ainsi que l'a démontré Broca, ni les caractères de l'os intermaxillaire, ni la structure des mains et des pieds,

(1) *Histoire de la création des êtres organisés, d'après les lois naturelles*, trad. franç. de Ch. Letourneau, p. 592 et 614.

ni la constitution et les fonctions de la colonne vertébrale, ni la conformation du bassin et du sternum, ni le système musculaire, ni les faits relatifs aux appareils sensoriaux externes, ni l'appareil digestif, ni les caractères anatomiques ou morphologiques du cerveau ne détachent l'homme des anthropoïdes (1). Bien plus, il existe sous ce rapport un intervalle tout autrement considérable entre les singes inférieurs et les anthropoïdes qu'entre ces derniers et l'homme (2).

L'on s'est rejeté alors sur des caractères soi-disant non physiques. Mais il s'est trouvé que les animaux inférieurs possédaient la prévoyance, la mémoire, l'imagination, le raisonnement, la pudicité, la dose de volonté compatible avec le déterminisme organique, et qu'ils donnaient les témoignages les moins équivoques de sentiments de pitié, d'admiration, d'ambition, d'affection, d'amour de la domination, d'initiative dans le travail.

En fin de compte, il fallut produire les deux arguments que l'on tenait en réserve ; l'argument de la *religiosité*, l'argument de la *moralité*. Leur succès fut malheureux.

Il est aisé, en effet, de soumettre la religiosité à la même critique dont relèvent toutes les manifestations intellectuelles et de démontrer que son origine n'est que la terreur, la crainte d'un inconnu : *Primus in orbe deos fecit timor.* L'enfant ne vient jamais au monde doué d'une faculté religieuse : « Il sait là-dessus ce qu'on lui en-

(1) Broca. *Discours sur l'homme et les animaux*, Bulletins de la Société d'anthropologie de Paris, 1866, p. 53. *L'ordre des primates. Parallèle anatomique de l'homme et des singes*, ibid., 1869, p. 228. *Etudes sur la constitution des vertèbres caudales chez les primates sans queue*, *Revue d'anthropologie*, t. II, p. 577. Consultez encore sur cet important sujet : Vogt. *Leçons sur l'homme, huitième leçon.* Schaaffhausen. *Les questions anthropologiques de notre temps.* Revue scientifique, 1868, p. 769. Paul Bert, Bulletins de la Société d'anthropologie de Paris, 1862, p. 473. Bertillon. *Ibid.*, 1865, p. 605.

(2) Broca. *L'ordre des primates*, etc., *op. cit.*, *passim*. Dally. *L'ordre des primates et le transformisme.* Bulletins de la Société d'anthropologie de Paris, 1868, p. 673.

seigne, mais il ne devine rien ; il n'en a pas la connaissance intuitive (1). » C'est ce que Broca a exposé en termes excellents : « L'auteur d'une conception religieuse met en jeu des facultés actives, parmi lesquelles l'imagination joue le principal rôle. Voilà une première espèce de religiosité active ; mais elle ne se manifeste que chez un très-petit nombre d'individus. La plupart, l'immense majorité des hommes, n'ont qu'une religiosité passive, qui consiste purement et simplement à croire ce qu'on leur dit sans avoir besoin de le comprendre, et cette religiosité n'est le plus souvent qu'un résultat de l'éducation. Dès l'âge le plus tendre, l'enfant est élevé au milieu de certaines croyances ; on y façonne son esprit sans qu'il soit en état de discuter et de raisonner. Aucune intelligence ne peut se soustraire à l'action de cet enseignement, combiné et perfectionné depuis des siècles. L'enfant s'y soumet toujours, et souvent d'une manière définitive. Il croit sans examen, parce qu'il n'est pas encore capable d'examiner, et parce que, pour toutes les notions, religieuses ou autres, il s'en rapporte aveuglément à l'autorité de ses instituteurs. Il n'y a rien dans tout cela qui puisse nous révéler l'existence d'une faculté, d'une aptitude ou d'une aspiration particulière. Mais avec l'âge, avec l'expérience, avec l'étude surtout, cet état passif de l'esprit fait place presque toujours à un certain degré de scepticisme. On apprend à se méfier plus ou moins de la parole d'autrui. Il ne suffit plus d'entendre dire une chose pour y croire ; on demande des preuves, et lorsqu'un individu accepte sans examen tout ce qu'on lui raconte,

(1) Letourneau. *De la religiosité et des religions au point de vue anthropologique*, Bulletins de la Société d'anthropologie de Paris, 1865, p. 581. *Sur la méthode qui a conduit à établir un règne humain*, ibid., 1866, p. 269. Lagneau. *Sur la religiosité*, ibid., 1865, p. 648. Coudereau. *Sur la religiosité comme caractéristique*, ibid., 1866, p. 329. Broca. *Discours sur l'homme et les animaux*, ibid., 1866, p. 59 et 74. Dally. *Du règne humain et de la religiosité*, ibid., 1866, p. 121.

on dit de lui qu'il est crédule comme un enfant. Cet esprit de critique, dont le développement marche de front avec celui de l'intelligence elle-même, s'applique d'abord aux notions matérielles, aux faits de la vie ordinaire, et souvent il ne s'étend pas au-delà de cet ordre de phénomènes; mais, souvent aussi, et sans changer de nature, il s'étend aux conceptions métaphysiques et religieuses ; de sorte que, dans tous les pays, surtout dans ceux où l'homme cultive son intelligence, on voit un grand nombre d'individus abandonner peu à peu une partie de la totalité de leurs croyances. Ce prétendu caractère humain, que vous appelez la religiosité, a donc disparu chez eux ? Les mettrez-vous au rang des brutes, ces hommes qui souvent se font remarquer par l'étendue de leur savoir, par la puissance de leur esprit ? Ainsi, de quelque manière qu'on envisage la religiosité, il est impossible de la considérer comme un fait général et inséparable de la nature de l'homme. La religiosité active, créatrice des conceptions religieuses, n'existe que chez de rares individus. La religiosité passive, qui n'est qu'une forme de la soumission à l'autorité, de l'appropriation d'une intelligence au milieu dans lequel elle se développe, est incomparablement plus répandue ; mais elle est bien loin d'être universelle; si elle l'était, les adeptes de toutes les religions ne tonneraient pas tant contre les incrédules. »

Il importe de bien le remarquer, non-seulement cette prétendue caractéristique arrive à faire défaut chez une grande part des hommes de science, mais encore elle manque absolument chez nombre de peuplades réputées sauvages. Nous n'avons que faire de reproduire ici les assertions fort catégoriques, et que l'on a vainement révoquées en doute, d'une foule d'observateurs désintéressés. L'on a prétendu que les peuples vivant sans dogmes et sans culte croyaient au moins à des forces et à des manifestations surnaturelles. Mais il est certain,

il est évident que l'infériorité même de ces peuples leur rend impossible toute distinction du naturel et du soi-disant surnaturel. Il en faut toujours revenir à cette terreur très-explicable dont nous parlions tout à l'heure, à la crainte d'un inconnu, ou, pour mieux dire de l'inconnu. S'il convient de voir là une croyance, il n'est point alors d'animal, même très inférieur, à qui l'on puisse contester la religiosité.

Nous ne voulons pas nous appesantir sur la dernière objection, la prétendue caractéristique tirée de la *moralité*. C'est un fait avéré qu'elle manque tout aussi bien chez beaucoup de peuples sauvages, comme nous l'enseigne l'ethnographie, et qu'on la rencontre évidente, éclatante, dans les actes d'un grand nombre d'animaux, au moins d'animaux sociables.

C'est la faculté du langage articulé qu'il faut invoquer, en définitive, pour distinguer l'homme de ses frères inférieurs. Chez aucun de ces derniers, en effet, l'on n'a pu rencontrer cette faculté. On n'a que faire d'arguer ici des paroles du perroquet, paroles articulées sans doute, mais dont l'émission est essentiellement distincte d'une conception corrélative ; il s'entend de soi que cette corrélation, cette connexité, est précisément la caractéristique du langage articulé humain ; le perroquet n'est qu'un écho inconscient.

Par contre, cette caractéristique du langage articulé est commune à toutes les races humaines. C'est là un fait concluant. Si baroque que nous puissent sembler les idiomes des dernières couches de l'humanité, ils n'en ont pas moins droit au nom de véritables langues, et leur plus ou moins d'harmonie et de charme n'a que faire en cette question. Notons que le plus souvent c'est leur matériel phonétique qui doit nous paraître étrange et non leur structure.

Mais, a-t-on dit, les individus qui ne donnent aucun

signe de cette prétendue caractéristique humaine, les sourds-muets de naissance, par exemple, ou les gens atteints d'aphasie par suite d'une lésion cérébrale, ne devraient pas, à ce compte, recevoir le nom d'*hommes*, et pourtant il est manifeste, il est incontestable que l'on ne peut point ne pas les tenir pour tels.

Cette double objection est à peine spécieuse. Il n'est pas inutile cependant de la réfuter.

Ce qui manque au sourd-muet de naissance, ce n'est en aucune façon la faculté dont il est ici question, c'est la liberté de mettre en action ladite faculté. Un sourd-muet n'est muet que parce qu'il est sourd ; c'est sa surdité qui entrave seule l'usage de la faculté du langage. Au surplus, un enseignement spécial peut rompre cette entrave, et le sourd-muet de naissance apprend à parler, apprend à se servir de la faculté native du langage articulé. Il existe des écoles particulières où on lui enseigne expérimentalement à proférer, au moyen du jeu de son appareil vocal, les sons que ses oreilles ne lui ont pas appris à connaître. « Le sourd-muet, en effet, étant l'individu qui n'est muet que par cela qu'il est sourd, l'individu qui ne parle pas, uniquement parce qu'il n'a pas entendu parler, l'organe qui fait défaut chez lui est celui de l'audition, et non celui du langage. Le sourd-muet proprement dit n'est pas plus atteint, dans les organes cérébraux de la parole, comme dans ses organes vocaux, que ne l'est, dans les organes de la locomotion, un individu auquel on a lié les jambes. Pas plus à l'un qu'à l'autre, la faculté native ne manque. Il ne leur manque à tous deux que la liberté de faire usage de cette faculté, et cela par suite d'un événement étranger à la faculté même (1). »

Nous nous arrêterons un peu plus longtemps sur le

(1) Vaïsse. Bulletins de la Société d'anthropologie de Paris. 1866, p. 146.

cas de l'abolition de la faculté du langage articulé, résultant d'une lésion cérébrale. Certes, il n'y a point de doute que les individus victimes d'une telle lésion ne conservent leur caractéristique naturelle, c'est-à-dire leur qualité d'homme, quand bien même l'aphasie, chez eux, se trouve complète ; mais le résultat des études importantes faites en France sur ce sujet ne nous semble pas assez connu, et il est bon, il est nécessaire de le répandre davantage. Cela, d'ailleurs, peut contribuer à bien mettre en relief la véritable nature des recherches linguistiques.

Les tentatives de localisation cérébrale entreprises au XVIIIe siècle partaient d'un principe sensé, mais le défaut de procédés d'expérience devait les faire avorter. Elles avortèrent en effet. De nos jours, l'anatomie pathologique a repris la question, et il est difficile de méconnaître la grande importance des résultats auxquels est arrivé Broca. Nous le suivrons d'une façon rapide.

L'exercice de la faculté du langage articulé est subordonné « à l'intégrité d'une partie très circonscrite des hémisphères cérébraux et plus spécialement de l'hémisphère gauche. Cette partie est située sur le bord supérieur de la scissure de Sylvius, vis-à-vis l'insula de Reil, et occupe la moitié postérieure, probablement même le tiers postérieur seulement de la troisième circonvolution frontale ».

C'est l'autopsie des aphasiques qui a démontré cette localisation. Dans cette autopsie, en effet, on découvre constamment « une lésion très-évidente de la moitié postérieure de la troisième circonvolution frontale gauche ou droite, » presque toujours, dix-neuf fois sur vingt, de la circonvolution du côté gauche. Une lésion grave de la circonvolution droite a souvent laissé persister l'usage de la parole, mais « l'on n'a jamais vu paraître la faculté du langage articulé chez les individus qui ont présenté à l'autopsie une lésion profonde des deux circonvolutions

en question » (1). Nous ne relaterons pas ici la série des observations, très-convaincantes à notre avis, recueillies à ce sujet par nombre d'anatomistes ; les lecteurs curieux de détails précis peuvent en chercher dans les ouvrages indiqués à la note précédente. Toutefois une question intéressante à soulever, c'est celle de savoir pour quel motif l'exercice de la faculté du langage articulé dépend d'une façon beaucoup plus particulière d'une circonvolution de l'hémisphère cérébral gauche, plutôt que de la circonvolution parallèle de l'hémispère droit, bien que les fonctions de l'un et de l'autre hémisphère ne semblent point être fondamentalement différentes. Ce fait curieux tient à ce que les circonvolutions de l'hémisphère gauche ont un développement en général plus rapide que celui des circonvolutions de l'hémisphère droit (2). Les premières se trouvent déjà dessinées, comme le dit Broca (3), à un moment où les autres ne sont pas encore apparentes. Il ajoute : L'hémisphère gauche, qui tient sous sa dépendance le mouvement des membres droits, est donc plus précoce dans son développement que l'hémisphère opposé. On comprend ainsi pourquoi, dès les premiers temps de la vie, le jeune enfant se sert de préférence des membres dont l'intervention est alors la plus parfaite, pourquoi, en d'autres termes, il devient droitier. Le membre supérieur droit, étant dès l'origine plus fort et plus adroit que le gauche, est appelé, par cela même, à fonctionner plus souvent, et il acquiert dès lors une supériorité de force et d'adresse qui ne fait que s'accroître avec l'âge. Jusqu'ici j'ai appelé *droitiers* ceux

(1) *Bulletins de la Société anatomique*, 1861, 1863. *Bulletins de la Société de Chirurgie*, 1864. *Bulletins de la Société d'anthropologie de Paris*, 1861, 1863, 1865, 1866. *Exposé des titres et travaux scientifiques*, 1868.

(2) Gratiolet, Bertillon, Baillarger.

(3) *Du siège de la faculté du langage articulé*, Bulletins de la Société d'anthropologie de Paris, 1865, p. 383.

qui se servent de préférence de la main droite, et *gauchers* ceux qui se servent de préférence de la main gauche. Ces expressions sont tirées de la manifestation extérieure du phénomène ; mais si nous considérons le phénomène par rapport au cerveau et non par rapport à ses agents mécaniques, nous dirons que la plupart des hommes sont naturellement gauchers du cerveau et que, par exception, quelques-uns d'entre eux, ceux qu'on appelle *gauchers*, sont au contraire droitiers du cerveau. Ce n'est ni dans les muscles, ni dans les nerfs moteurs, ni dans les organes cérébraux moteurs, tels que les couches optiques ou les corps striés, que gît le phénomène essentiel du langage articulé. Si l'on n'avait rien de plus que ces organes, on ne parlerait pas. Ils existent quelquefois, parfaitement sains et parfaitement conformés, chez des individus devenus complétement aphémiques ou chez des idiots qui n'ont jamais pu ni apprendre ni comprendre aucun langage. Le langage articulé dépend donc de la partie de l'encéphale qui est affectée aux phénomènes intellectuels, et dont les organes cérébraux moteurs ne sont en quelque sorte que les ministres. Or, cette fonction de l'ordre intellectuel, qui domine la partie dynamique aussi bien que la partie mécanique de l'articulation, paraît être l'apanage à peu près constant des circonvolutions de l'hémisphère gauche, puisque les lésions qui produisent l'aphémie occupent à peu près constamment cet hémisphère. Cela revient à dire que, pour le langage... nous sommes gauchers du cerveau... nous parlons avec l'hémisphère gauche. C'est une habitude que nous prenons dès notre première enfance. De toutes les choses que nous sommes obligés d'apprendre, le langage articulé est peut-être la plus difficile. Nos autres facultés, nos autres actions existent au moins à l'état rudimentaire chez les animaux ; mais quoique ceux-ci aient certainement des idées, et quoiqu'ils sa-

chent se les communiquer par un véritable langage, le langage articulé est au-dessus de leur portée. C'est cette chose complexe et difficile que l'enfant doit apprendre à l'âge le plus tendre, et il y parvient à la suite de longs tâtonnements et d'un travail cérébral de l'ordre le plus compliqué. Eh bien, ce travail cérébral, on le lui impose à une époque très rapprochée de ces périodes embryonnaires où le développement de l'hémisphère gauche est en avance sur celui de l'hémisphère droit. Dès lors, il ne répugne pas d'admettre que l'hémisphère cérébral le plus développé et le plus précoce soit, plus tôt que l'autre, en état de diriger l'exécution et la coordination des actes à la fois intellectuels et musculaires qui constituent le langage articulé. Ainsi naît l'habitude de parler avec l'hémisphère gauche, et cette habitude finit par faire si bien partie de notre nature, que, lorsque nous sommes privés des fonctions de cet hémisphère, nous perdons la faculté de nous faire comprendre par la parole. Cela ne veut pas dire que l'hémisphère gauche soit le siège exclusif de la faculté générale du langage, qui consiste à établir une relation déterminée entre une idée et un signe, ni même de la faculté spéciale du langage articulé, qui consiste à établir une relation entre une idée et un mot articulé ; l'hémisphère droit n'est pas plus étranger que le gauche à cette faculté spéciale, et ce qui le prouve, c'est que l'individu rendu aphémique par une lésion profonde et étendue de l'hémisphère gauche, n'est privé en général que de la faculté de reproduire lui-même les sons articulés du langage ; il continue à comprendre ce qu'on lui dit et, par conséquent, il connaît parfaitement les rapports des idées avec les mots. En d'autres termes, la faculté de concevoir ces rapports appartient à la fois aux deux hémisphères, qui peuvent, en cas de maladie, se suppléer réciproquement ; mais la faculté de les exprimer par des mouvements coordonnés dont la pratique ne s'ac-

quiert qu'à la suite d'une très-longue habitude, paraît n'appartenir qu'à un seul hémisphère, qui est presque toujours l'hémisphère gauche. Maintenant, de même qu'il y a des individus gauchers, chez lesquels la prééminence native des forces motrices de l'hémisphère droit donne une prééminence naturelle et incorrigible aux fonctions de la main gauche, de même on conçoit qu'il puisse y avoir un certain nombre d'individus chez lesquels la prééminence native des circonvolutions de l'hémisphère droit renversera l'ordre des phénomènes que je viens d'indiquer ; chez lesquels, dès lors, la faculté de coordonner les mouvements du langage articulé deviendra, par suite d'une habitude contractée dès la première enfance, l'apanage définitif de l'hémisphère droit. Ces individus exceptionnels seront, par rapport au langage, comparables à ce que sont les gauchers par rapport aux fonctions de la main. Les uns et les autres seront droitiers du cerveau... L'existence d'un petit nombre d'individus qui, par exception, parleraient avec l'hémisphère droit expliquerait très bien les cas exceptionnels où l'aphémie est la conséquence d'une lésion de cet hémisphère. Il suit de ce qui précède qu'un sujet chez lequel la troisième circonvolution frontale gauche, siège ordinaire du langage articulé, serait atrophiée depuis la naissance, apprendrait à parler et parlerait avec la troisième circonvolution frontale droite, comme l'enfant venu au monde sans la main droite devient aussi habile avec la main gauche qu'on l'est ordinairement avec l'autre main (1). »

Nous n'avons qu'un mot à ajouter à cette citation, c'est que les observations recueillies jusqu'à ce jour et dont le nombre est maintenant considérable, viennent toutes

(1) Consultez également Adr. Proust, *Altérations de la parole*, Bulletins de la Société d'anthropologie de Paris, 1873, p. 786. Du même auteur : *De l'aphasie*, Archives générales de médecine. Paris, 1872.

confirmer la doctrine de cette localisation de la faculté du langage articulé.

C'est là un fait capital et qui en dit à lui seul plus que tous les autres, lorsqu'il s'agit de démontrer que l'étude du langage articulé est du domaine de l'histoire naturelle, ainsi que nous avons déjà cherché à établir dans le chapitre précédent.

La possession de la faculté du langage articulé ne présage rien d'ailleurs de ce que sera, chez l'individu qui s'en trouve doué, l'exercice de cette faculté. Cet exercice en effet est un art, un art difficile : l'enfant bégaye et bégaye longtemps, jusqu'au jour où, grâce à un certain développement intellectuel, grâce également à l'habitude acquise, il parvient à user comme ceux qui l'entourent de sa faculté native. En d'autres termes, la faculté est naturelle, mais l'usage de cette faculté est un art : la première a été assez heureusement qualifiée, en grec, de ἐνέργεια ; le second, de ἔργον.

De là les actes purement automatiques qui se révèlent en si grand nombre dans l'exercice de la fonction dont il s'agit, tant dans ses manifestations normales qu'à l'état pathologique (1).

Cette distinction est importante, et l'on risquerait, en la négligeant, de se former sur l'origine du langage les conceptions les plus bizarres et les moins scientifiques.

Hérodote raconte, au livre second de ses *Histoires*, que Psammétique, roi d'Egypte, voulant connaître quel était le plus ancien des peuples, confia à un pâtre deux enfants nouveau-nés ; ceux-ci devaient vivre dans l'isolement et n'entendre aucune voix humaine. Des chèvres pourvoyaient à leur allaitement. Au bout de deux ans, le pâtre fut reçu par ces enfants au cri répété de βέκος.

(1) Onimus, *Bull. de la Soc. d'anthr.*, 1873, p. 759. Ferrier, *Localisation des maladies cérébrales*, 1880.

Psammétique, après enquête, découvrit que ce mot de βέκος appartenait à la langue phrygienne et qu'il voulait dire « pain ». Les Egyptiens durent reconnaître alors que leur origine était moins ancienne que celle des Phrygiens.

Dans ce conte ridicule nous voyons deux enfants inventer, sans connaître aucun autre mot, un nom incontestablement dérivé, et selon toute vraisemblance, décliné. C'est bien là un exemple de la critique des anciens. Admettons que l'expérience dont il s'agit ait eu lieu réellement, est-ce bien le mot βέκος que ces enfants ont prononcé ? N'ont-ils pas imité plutôt, et tout simplement, la voix de leur nourrice ?

En tous cas, l'idée de Psammétique dénote la complète ignorance de ce fait capital et hors de toute discussion que l'exercice du langage articulé est un art difficile, un art acquis et que les générations se sont transmis les unes aux autres. Comment attendre d'un individu en présence duquel on n'aura jamais ouvert la bouche, qu'il connaisse et parle une langue quelconque ? Une langue ne s'invente pas ; une langue toute faite, le phrygien comme toutes les autres, a déjà parcouru plusieurs périodes de sa vie. Ici, comme en toutes choses, le présent est la résultante du passé. Comment un individu isolé pourrait-il à lui tout seul créer à nouveau cette longue série de phases diverses qu'ont connues toutes les langues ? L'on ne fabrique pas un système linguistique ; il se forme et se développe de lui-même, par degrés, petit à petit, mais il est né en même temps qu'est né l'homme : non pas l'homme individu, mais l'homme pris dans le sens général, le groupe humain, si l'on veut. C'est ce que nous avons dit plus haut : l'apparition de la faculté du langage articulé détermine le point d'évolution où un primate a droit au nom d'*homme*.

Schleicher, dans sa rapide mais si substantielle notice

sur l'importance du langage pour l'histoire naturelle de l'homme et dans son écrit non moins remarquable sur la théorie darwinienne et la science du langage, a traité de cette corrélation de la naissance de l'homme et de l'apparition du langage articulé. « Si c'est le langage qui fait l'homme, dit-il, nos premiers pères n'ont pas été réellement hommes : ils ne le sont devenus qu'au moment où se forma le langage, et cela grâce au développement du cerveau, grâce au développement des organes de la parole. » La linguistique, comme toutes les autres sciences naturelles, nous force à admettre que l'homme s'est développé de formes inférieures ; qu'il est *devenu* homme, mais qu'il n'est pas *né* homme par un coup de baguette quelconque.

Nous avons à notre tour repris ce sujet, lors de l'excellente communication sur le *Précurseur de l'homme*, faite par G. de Mortillet à l'Association française pour l'avancement des sciences (1), au sujet des silex taillés trouvés dans les couches marneuses de l'étage des calcaires de Beauce. D'après les lois de la paléontologie, l'homme actuel ne devait pas exister à cette époque ; la succession des faunes dans les divers étages géologiques est en effet reconnue et acquise : d'étage en étage les animaux se modifient, et leurs variations se précipitent d'autant plus que leur organisation est plus compliquée. Trois fois au moins la faune s'est renouvelée depuis l'époque de formation du calcaire de Beauce, et les mammifères du niveau des marnes à silex dont il est question appartiennent à des germes éteints, à des genres prédécesseurs mais distincts des genres aujourd'hui vivants. On ne peut admettre avec quelque raison que l'homme seul ait échappé à cette variation, l'homme, précisément, dont l'organisation est des plus compliquées : la taille des

(1) Seconde session, tenue à Lyon, août 1873.

silex de l'époque tertiaire moyenne serait donc due à un genre précurseur de l'homme. Cette opinion revêt, à nos yeux, les caractères de la plus haute vraisemblance, et elle répond de tous points à la doctrine exposée par Schleicher dans les opuscules dont nous parlions ci-dessus.

Si nous ne pouvons admettre, sans tomber dans des conceptions métaphysiques et sans fondement, que la faculté du langage articulé ait été un beau jour acquise à l'homme sans cause, sans origine, *ex nihilo*, il nous faut bien accepter alors qu'elle est le fruit d'un développement progressif des organes. Cela suppose avant l'homme, avant l'être caractérisé par la faculté du langage articulé, un autre être en train d'acquérir cette faculté, c'est-à-dire en voie de devenir homme. Ainsi que l'enseigna Schleicher, il faut admettre qu'un certain nombre seulement de ces êtres encore dépourvus de la faculté du langage articulé mais bien près de l'acquérir la gagnèrent en réalité, sous l'influence des conditions heureuses, et dès lors eurent réellement droit à la dénomination d'*hommes ;* mais que, par contre, un certain nombre d'entre eux, moins favorisés par les circonstances, échouèrent dans leur développement et tombèrent dans la métamorphose régressive : nous aurions à reconnaître leurs restes dans les anthropomorphes, gorilles, chimpanzés, orangs, gibbons. Nous verrons plus loin, lorsqu'il s'agira de passer l'examen des différentes couches du langage, que ces couches diverses témoignent de la façon la moins équivoque d'un progrès constant, d'un développement naturel, d'un perfectionnement régulier.

D'ailleurs, en présence de ce perpétuel spectacle d'évolution qui se déroule sous nos yeux dans la nature entière, nous ne pouvons pas ne pas admettre que la faculté du langage articulé ne se soit acquise petit à petit, grâce à un développement progressif des organes.

Et peu importe que ce développement soit dû aux différentes sortes de sélection ou qu'il provienne d'autres causes, inconnues encore à ce jour. C'est un sujet sur lequel nous ne pouvons nous étendre ; il appartient à l'étude générale de la variabilité et de la transformation des espèces, et nous devons nous contenter de l'indiquer. Ici, sans doute, comme partout ailleurs, la fonction a été pour beaucoup dans les progrès de l'organe lui-même, mais ici également, comme partout, l'organe tel qu'il est, l'organe sous sa forme actuelle, n'a pu que procéder d'une forme inférieure.

Il faut donc reconnaître, en définitive, que cette caractéristique de l'homme, la faculté du langage articulé, est purement relative. Nous découvrons son origine et ses rudiments (1) ; nous comprenons que nos pères ne l'ont acquise que par degrés, dans le combat pour le progrès d'où ils devaient sortir victorieux.

Mais, pour être relative, cette faculté n'en est pas moins particulière, spéciale à l'homme, et, au demeurant, c'est grâce à elle seule que le premier des primates peut porter ce nom d'*homme* qu'il a gagné, à travers des milliers de siècles, au prix de luttes incessantes.

(1) Lamarck, *Philosophie zoologique*, t. I. — Darwin, *La Descendance de l'homme*, traduct. franç., t. I, p. 59. — Hæckel, *Histoire de la création des êtres organisés*, traduct. franc., p. 591. — Whitney, p. 234.

CHAPITRE III.

PREMIÈRE FORME LINGUISTIQUE : LE MONOSYLLABISME LES LANGUES ISOLANTES.

Parmi les formes différentes que peuvent présenter les langues ou les familles de langues, la forme monosyllabique est la plus simple ; c'est la forme élémentaire, chez laquelle les mots sont de simples racines. Ces racines-mots, ou ces mots-racines, n'éveillent qu'une idée essentiellement générale. Nulle indication de personne, de genre, de nombre ; nulle indication de temps, de mode ; point d'éléments de relation, point de conjonctions, point de prépositions. Rien qu'une idée très large, sinon très vague, une idée que ne rend même pas la forme, si peu déterminée déjà, de notre infinitif.

Dans ce premier état (nous dirons plus tard dans cette première couche), la forme du mot est donc unique : c'est la racine telle quelle, la racine invariable. La langue, dans cette première étape, n'est formée que d'éléments dont le sens est éminemment général : point de suffixes, point de préfixes, aucune modification, quelle qu'elle soit, qui puisse indiquer une relation, un rapport quelconque. A ce premier degré, le plus simple de tous, la phrase est donc faite d'après cette formule : *racine* + *racine* + *racine*, etc., etc., et ces racines successives (c'est là le point capital à noter) sont toujours invariables.

On comprend, après ce court exposé, pourquoi les langues de cette espèce ont reçu la dénomination générale de *monosyllabiques* ou d'*isolantes :* leurs mots en

effet sont formés de simples racines monosyllabiques, isolées, indépendantes en principe les unes des autres.

Il est bon de le dire dès maintenant, tous les systèmes linguistiques ont passé par cette période du monosyllabisme ; les langues les plus complexes sous le rapport de la forme, c'est-à-dire les langues à flexion — telles, par exemple, que les langues indo-européennes — révèlent à l'analyse scientifique les traces non équivoques d'une origine monosyllabique, origine lointaine et à laquelle elles ne remontent que par l'intermédiaire d'un autre état, mais que l'on ne saurait mettre en doute un seul instant. C'est ce que nous aurons à constater en temps opportun. Nous verrons aussi, au moment voulu, que la forme intermédiaire, la période de l'agglutination — où l'on rencontre, par exemple, le basque, le japonais, les langues dravidiennes — a donné naissance au système de la flexion, mais qu'elle provient, elle-même, de la couche inférieure, celle du monosyllabisme qui nous occupe en ce moment.

Ce n'est pas à dire que toutes les langues agglutinantes doivent se changer quelque jour en langues à flexion, ni que toutes les langues isolantes (c'est-à-dire monosyllabiques) soient appelées à devenir agglutinantes. Non, sans doute. Bien des langues ont péri qui appartenaient aux deux classes inférieures, et il est certain que, parmi les langues aujourd'hui vivantes et qui se trouvent soit à l'étage du monosyllabisme, soit à celui de l'agglutination, le plus grand nombre est fixé d'une manière définitive ; l'on peut dire, par exemple, sans hésitation que le basque, que les idiomes des Indiens de l'Amérique septentrionale périront sous leur forme actuelle.

D'ailleurs, ce n'est pas sans causes déterminantes que telle ou telle langue s'est fixée de façon définitive dans telle ou telle couche, par exemple, dans celle du monosyllabisme ou dans celle de l'agglutination, et

qu'elle ne manifeste plus que des tendances très faibles et très rares à atteindre la couche supérieure. Il se peut que ces motifs aient été multiples, qu'ils aient été d'ordre fort divers, et le soin de les découvrir est une tâche ardue.

Cette tâche n'a pas encore été abordée. Elle doit avoir pourtant un heureux succès. Il y a motif à tout, et chaque jour on fait un pas du connu à l'inconnu.

Nul doute, au surplus, que la plus puissante de ces causes n'ait été l'entrée dans la vie historique et la production littéraire. Cette production témoigne déjà par elle-même, par elle seule, que la langue se suffisait telle quelle et se sentait en état, s'il est permis de s'exprimer ainsi, de répondre à tous les besoins d'une nation constituée. En ce sens, il n'est pas inexact de dire qu'à son premier pas dans la vie historique, l'homme atteint la période que l'on appelle en histoire naturelle la *période de métamorphose régressive*. C'est ce que l'avenir confirmera ou infirmera ; mais il n'est guère possible, à l'heure présente et dans les conditions scientifiques actuelles, de n'émettre que des assertions plus ou moins conjecturales.

Il est aisé de comprendre que le système d'une succession de racines, à idées toujours très générales, ne devait offrir au langage que des moyens fort restreints. Il est impossible que le besoin inévitable d'exprimer les rapports ne se soit pas fait sentir de très bonne heure ; or, ainsi que nous l'avons dit, la succession de mots-racines, ou, pour parler de façon plus exacte, de racines-mots, était la négation, l'exclusion même des éléments de relation, des éléments appelés à n'indiquer que les rapports : rapports d'activité ou de passivité, d'unité ou de pluralité de passé, de présent, de futur. Une telle période, cependant, a dû exister. Il la faut reléguer, sans aucun doute, en des âges préhistoriques très lointains, et, selon toute vraisemblance, elle succéda à l'âge plus ancien

encore durant lequel se constituèrent les racines par le fait de l'agrégation des éléments simples phoniques.

L'on remédia par un expédient ingénieux à ce défaut de détermination. Ce fut en réglant d'une façon très rigoureuse la place que devaient occuper les racines, c'est-à-dire les mots dans l'ensemble de la phrase.

La syntaxe était née ainsi avant la grammaire proprement dite. Comme nous aurons à le constater ce procédé de la position forcée des mots donna naissance par la suite à la seconde forme linguistique, celle de l'agglutination. En jetant un coup d'œil rapide sur les diverses langues monosyllabiques, nous verrons comment on usa de cette ressource importante et comment aussi son origine put s'obscurcir peu à peu.

Quoi qu'il en soit, l'on voit déjà que la grammaire de toute langue monosyllabique, c'est-à-dire de toute langue isolante, est et ne peut être qu'une syntaxe. Dans ces langues en effet le mot est inflexible ; en dépit de tout changement de position dans la phrase, il demeure invariable, toujours le même, et c'est uniquement la position qu'il occupe qui détermine sa valeur, sa qualité de sujet ou de régime, d'épithète ou de substantif, de verbe ou de nom, et ainsi de suite.

Il faut remarquer encore, d'une façon générale, que l'importance de l'intonation est considérable dans les langues monosyllabiques ; ce point ne nous semble pas avoir été traité d'une manière assez complète dans les différents écrits sur les langues en question. La grande valeur du ton, de l'intonation, n'est pas de différencier à l'occasion un grand nombre d'homophonies, c'est-à-dire de mots identiques quant à la forme, mais divers quant à leur signification respective.

Le *chinois*, l'*annamite*, le *siamois*, le *birman*, le *tibétain* sont les langues monosyllabiques principales. Ils constituent ou représentent tout autant de systèmes glotti-

ques indépendants les uns des autres et que l'on ne pourrait ramener à une origine commune.

Il existe d'ailleurs d'autres langues monosyllabiques dans la péninsule indo-chinoise, telles que le *pégou* dans la Birmanie anglaise, et le *kassia* dans une petite région située à deux cents milles anglais du fond de la mer du Bengale, sur la rive gauche du Brahmapoutra, au sud de l'Assam. Leur peu d'importance nous autorise à les passer sous silence.

Nous n'avons ici ni le dessein ni la possibilité de passer à tour de rôle en revue ces différentes langues ; nous nous contenterons de donner sur chacune d'elles quelques renseignements généraux, en insistant davantage sur la langue chinoise, la plus caractéristique de toutes les langues de cette espèce.

§ 1. Le chinois.

Les trois grand dialectes du chinois sont : la langue mandarine (vulgaire dans les provinces centrales et usitée, en tant qu'idiome cultivé, dans tout l'empire) ; le dialecte de Canton ; le dialecte de Foukian. Tous trois, d'ailleurs, pour appartenir à la même langue, sont profondément distincts, et il est bien difficile que les habitants du Nord et ceux du Sud se comprennent les uns les autres.

L'étude du chinois se compose de deux parts nettement tranchées : l'écriture, la langue elle-même.

Parlons en premier lieu de cette dernière.

Ainsi que nous l'avons dit, elle est purement et simplement syntaxique. Le premier écueil qu'il lui fallut éviter fut, comme pour toutes les langues isolantes, l'indécision très fréquente du sens, étant donnée la multiplicité des significations que peut revêtir chez elle une seule et même forme. La forme *tao*, par exemple, signifie

indistinctement (et entre autres acceptions), *ravir, atteindre, couvrir, drapeau, froment, mener, chemin ;* la forme *lu* (entre autres acceptions également) signifie *détourner, véhicule, pierre précieuse, rosée, forger, chemin.* Ce fut un procédé naïf, simple, mais très-exact, que de faire se succéder deux termes capables d'être synonymes en l'une quelconque de leurs acceptions : par exemple, *tao* et *lu*, qui répondent l'un et l'autre à l'idée de *chemin.* Ce procédé fut employé : *tao* laisse le choix entre neuf ou dix sens, mais *tao lu* ne peut dire que *chemin.* Est-ce là, comme on l'a prétendu, une véritable *composition,* la fabrication d'un vrai *composé ?* En aucune façon ; un composé indique toujours une relation, et ici il n'y a qu'une accumulation de synonymes.

On ne peut voir non plus des composés réels — bien qu'il en puisse sembler au premier abord — dans l'association des mots *fu* « père » et *mu* « mère » qui signifient « parents », de *yuan* « éloigné » et *kin* « près » qui signifie « distance ». En effet, dans ces accumulations de synonymes, le premier mot ne dépend pas du second, le second ne dépend pas du premier.

Le genre d'un mot ne peut être déterminé, on le conçoit, qu'à l'aide d'un second terme. On a recours, par exemple, à *nan* « mâle, masculin », *niu* « femelle, féminin » ; de là : *nan tse* « fils », *niu tse* « fille », *niu jin* « femme ». S'agit-il d'animaux, les termes sont différents, mais le procédé reste le même. Il est assurément des plus simples : nous le retrouverons plus loin dans les langues agglutinantes, en wolof, en japonais, et plus tard encore dans les idiomes les plus développés. En latin, par exemple, nous rencontrerons *mas canis, femina canis, femina porcus, anguis femina* et bien d'autres expressions analogues. Combien de phénomènes appartenant en propre à la première phase linguistique ont persisté à travers les âges jusqu'à la dernière période !

Singulier ou pluriel, le nombre n'est indiqué, en principe, que par l'ensemble même de la phrase. Parfois, cependant, on emploie un terme dont les sens est celui de *multitude, de totalité : to jin*, une foule de gens, beaucoup de gens, « les gens ».

Le sujet s'indique de lui-même, par ce fait qu'il commence toujours la proposition. Le régime direct, si la phrase est simple, se révèle aussi de lui seul en ce qu'il prend place immédiatement après le terme désignant l'action ; c'est le procédé que nous appliquerions en disant : « Emile craint Auguste, » et « Auguste craint Emile ». Mais, en d'autres circonstances, c'est l'emploi nécessaire de certains mots qui détermine le régime direct. Ces mots auxiliaires peut-on les regarder comme de véritables propositions ? Non certes, en aucun cas. Ce ne sont toujours que des racines-mots, car le chinois ne connaît point d'autres termes, ainsi que nous l'avons dit. Mais que ces racines, que ces mots auxquels ont fait ainsi appel, conservent encore et toujours dans l'esprit de ceux qui les emploient leur propre et indépendante valeur, c'est ce que l'on ne saurait admettre. Cette valeur s'atténue peu à peu, elle se subordonne, et cette subordination même est la cause qui des langues isolantes fait, avec le temps, des langues monosyllabiques.

La notion du locatif, celle du datif, celle de l'instrumental, celle de l'ablatif sont également rendues soit par l'accession de certains mots, soit par la place dans la phrase. Il suffit d'indiquer ce fait en général, sans entrer dans l'exposition d'une série d'exemples qui nous déborderaient et qu'il est facile de trouver dans les ouvrages spéciaux. Quant au génitif, on l'exprime clairement en plaçant le terme principal après le terme relatif: *thien tse* « fils du ciel » ; ou bien encore, on introduit entre ces deux mots ainsi placés le terme *ti* (en langue mandarine).

C'est par des procédés tout analogues que l'on rend la notion de qualification et celle de comparaison.

Enfin l'idée du verbe, sur laquelle repose la proposition tout entière, s'exprime encore d'une façon purement syntaxique, ou bien doit se déduire du sens général de la phrase. Rien, par exemple, n'indique en chinois la notion de notre temps imparfait ; parfois également on ne peut comprendre que par le sens général de la phrase qu'il s'agit de l'idée du futur.

Si nous passons de la notion du temps à celle de la modalité, au mode, nous constatons encore que c'est la position syntaxique qui indique le conditionnel. Quand au subjonctif et à l'optatif, ils se trouvent désignés par l'emploi de mots auxiliaires.

Ainsi, en chinois, il ne peut pas plus y avoir de verbe qu'il ne peut y avoir de nom. Nous ne saurions trop le répéter, c'est la syntaxe qui particularise le sens des mots et qui constitue toute la grammaire. En dehors de sa place dans la phrase, le mot n'est qu'une racine à acception aussi large que possible ; et c'est seulement quand il prend position qu'il éveille une idée d'individualité, de qualité, de relation, d'activité, une idée particularisée. C'est ainsi, par exemple, qu'une seule et unique forme *ngan* signifie « procurer le repos, jouir du repos, posément, repos »); une autre forme, *ta*,« grand, grandement, grandeur, agrandir » ; une autre forme, « rond, boule, en rond, arrondir » ; une autre forme encore, « être, vraiment, il, celui-ci, ainsi ».

Nous l'avons dit ci-dessus, et nous devons y revenir en temps opportun, l'emploi des mots accessoires, appelés à donner aux mots principaux le sens bien déterminé qui leur manque, fait passer les langues isolantes à l'état de langues agglutinantes. Le sens de ces racines accessoires s'est obscurci peu à peu ; on est venu, avec le temps, à ne plus leur accorder qu'une sorte de valeur un

peu arbitraire ; mais il fut une époque, une époque lointaine, l'âge d'or du monosyllabisme, pour ainsi dire, où leur sens véritable, leur signification pleine et entière, s'offrait seule et d'elle-même à l'esprit.

C'est un fait que les Chinois ont remarqué avec une sagacité surprenante, lorsqu'ils classèrent les racines en deux groupes distincts, les mots *pleins* et les mots *vides*. Par les premiers, par les mots pleins, ils entendaient les racines dont la signification restait dans toute sa plénitude et son indépendance, les racines que nous rendons dans nos traductions par des noms ou des verbes ; ils appelaient *mots vides* les racines dont la valeur propre s'obcurcissait par degrés et qui peu à peu recevaient la mission de déterminer et de préciser la notion très vague des mots pleins, des mots dont le sens primitif persistait tout entier. Observation remarquable et qui témoigne, mieux que bien d'autres découvertes, d'un esprit singulièrement perspicace. « Qu'est-ce que la grammaire ? » demande à son élève l'instituteur chinois. « C'est un art très utile, répond l'élève, un art qui nous enseigne à distinguer les mots pleins et les mots vides. »

Après avoir parlé de l'importance de la place syntaxique des racines et de leur valeur respective, il y a lieu de dire quelques mots des différentes inflexions de la voix en chinois.

Les différents *tons* que l'on rencontre en petit nombre dans la langue chinoise ont une utilité capitale lorsqu'il s'agit de distinguer les significations, parfois très diverses, de syllabes formées des mêmes éléments. Le vocabulaire chinois quasi académique donne quarante-deux mille caractères graphiques différents, ayant chacun leur prononciation propre ; or, comme la langue parlée ne possède environ que douze cents consonnances, « il faut donc que la même prononciation soit attachée en moyenne à plus de trente caractères » (d'Hervey Saint-

Denys). On voit que si l'intonation n'a pu venir à bout de toute difficulté, elle avait du moins une utilité bien considérable. Ce fait, nous l'avons dit, est commun aux diverses langues monosyllabiques. Les ouvrages spéciaux citent nombre d'exemples que nous n'aurions que faire de relater ici, et, sans entrer en plus de détails, nous n'avons qu'à mentionner ce procédé ingénieux et fort pratique.

Le matériel phonétique des Chinois n'est pas des plus complexes, mais on ne peut cependant le mettre au rang des plus simples. Parmi les consonnes, nous ne rencontrons ni *g*, ni *d*, ni *d* dans le dialecte mandarin ; dans le dialecte de Fukian, le *d* seul fait défaut. Dans ce dernier dialecte, les sifflantes sont moins variées que dans le précédent. L'absence de la consonne *r* est un fait bien connu. Les voyelles n'offrent rien de particulier ; on les rencontre souvent à l'état de diphthongues, et souvent aussi elles sont nasalisées.

En tout cas, et ceci est un fait caractéristique, le monosyllabe chinois s'ouvre par une consonne et se termine par une voyelle. Les signes *n* ou *ng*, que nous rencontrons à la fin des mots chinois transcrits en caractères latins, indiquent seulement la nasalisation des voyelles précédentes. Il n'est qu'un mot, un seul, qui échappe à cette règle sévère d'une consonne initiale et d'une voyelle terminale : *eul*, « deux » et « oreille ».

Les questions de graphique pure sortent du domaine de la linguistique ; elles constituent une étude spéciale, sans doute pleine d'intérêt, mais tout à fait distincte et indépendante. Il est utile pourtant de dire ici quelques mots du système graphique des Chinois et de montrer avec quelle habileté ce peuple sut appliquer à sa langue, si curieuse, un ensemble de caractères peu faits en apparence pour répondre à ce qu'on allait lui demander.

Etant donné le grand nombre d'homophonies d'une

langue monosyllabique, c'est-à-dire le grand nombre de syllabes formées des mêmes éléments phoniques, bien que répondant à des idées forts distinctes, il y avait une difficulté sérieuse à déterminer dans une système graphique les sens multiples des homophonies en question. Les Chinois arrivèrent à ce résultat par l'emploi de deux sortes de signes.

Leur première espèce de caractères ne se compose que d'images, que de vrais dessins : l'image d'un arbre, d'une montagne, d'un chien. Tantôt on les emploie indépendants, isolés ; tantôt on les accouple pour rendre une idée plus ou moins complexe. C'est ainsi que l'image de l'eau et celle d'un œil, si elles sont juxtaposées, rendent l'idée de *larmes ;* une porte et une oreille rendent l'idée *d'entendre ;* le soleil et la lune rendent l'idée d'*éclat.* Il faut également ranger parmi les véritables dessins les groupements de lignes ou de points, qui figurent, ou bien des nombres — un, deux, trois — ou bien l'état de supériorité, d'infériorité, d'inclinaison vers tel ou tel côté, et ainsi de suite.

Il fut un temps où ces caractères, où ces images, éveillaient d'une façon directe grâce à l'exactitude de leur représentation, la notion qu'ils étaient appelés à rendre. Mais peu à peu ces traits naïfs et véridiques perdirent leur forme originelle. Dans les signes qui laissent entendre aujourd'hui les idées de chien, de soleil, de lune, de montagne, on ne retrouve plus de prime abord les images anciennes qui évoquaient de façon directe ces diverses idées. Les caractères de cette première espèce ont été évalués au nombre minime d'environ deux cents (1).

La seconde sorte de caractères est plus compliquée.

(1) Abel Rémusat, *Recherches sur l'origine et la formation de la langue chinoise*, Mémoires de l'Académie des inscriptions et belles-lettres, 1820.

Elle comporte deux éléments : un élément phonétique et un élément idéographique.

Ainsi qu'on le comprend sans peine d'après tout ce qui a été dit ci-dessus, ce dernier élément a pour mission de déterminer la valeur parfois très multiple de l'élément phonétique. Ce dernier, si l'on ne figure que lui seul, laisse flotter l'esprit du lecteur entre un grand nombre d'homophones ; mais qu'on lui adjoigne un élément idéographique et l'hésitation cesse tout de suite : on a évoqué une idée déterminée, ou du moins une catégorie d'idées. C'est là un procédé fort ingénieux.

En somme, le caractère pris dans son ensemble, dans sa totalité, indique tout à la fois la prononciation et le sens. Ses deux parties se complètent réciproquement ; mais l'une de ces parties est regardée comme nulle quant à sa valeur phonique, et c'est l'autre qui détermine seule la prononciation. Si, par exemple, le signe *tcheu*, vaisseau, est accolé au-devant des signes qui représentent *huo*, feu, *ma*, cheval, ces deux derniers signes perdront leur valeur phonétique, le mot sera lu *tcheu*, mais ce *tcheu* ne signifiera plus vaisseau. Grâce au caractère dont il se trouve précédé, il laissera entendre soit un vacillement de la flamme, soit une sorte particulière de chevaux (1).

Les Chinois ont arrêté à 214 le nombre des signes, des caractères qu'ils ont appelés *chefs de classe*, et auxquels nous donnons le nom de *clefs*. Ces caractères comprennent, outre les 169 signes idéographiques (dont nous avons ci-dessus expliqué le rôle alors qu'ils se trouvent joints à un élément qui n'est que phonétique), une petite

(1) La grammaire chinoise de Stephan Endlicher est la plus simple de toutes celles que nous avons étudiées, mais l'absence de critique s'y fait trop souvent sentir. *Afangsgründe der chinesischen Grammatik*, Vienne, 1845. On étudiera avec profit les règles de position des mots dans la *Syntaxe nouvelle de la langue chinoise* de Stanislas Julien. Paris, 1869.

série de signes purement graphiques ou de simples images. Ces 214 clefs contiennent les éléments de tous les caractères chinois ; il y en a environ 50.000, dont 15.000 à peu près peuvent être en usage. A ces 214 clefs, il faut donc subordonner tous les autres caractères. C'est ce qu'ont fait les Chinois dans leur classification lexique, en ayant soin de disposer les clefs en un ordre consécutif, selon qu'elles se trouvaient représentées par un, deux, trois traits, et ainsi de suite ; la dernière en a dix-sept.

Cette classification arbitraire n'a rien à faire, ainsi qu'on le voit, avec la langue elle-même, et en effet nous avons dit plus haut que l'étude du chinois comprenait deux parts bien distinctes : celle de la langue, celle de l'écriture ; de là les difficultés très-sérieuses que rencontrent les commençants dans l'étude du chinois.

Ajoutons que tous les caractères peuvent être employés, en certaines occasions, comme s'ils n'étaient que phonétiques. C'est de cette façon que les Chinois peuvent écrire avec leurs signes des noms d'emprunt, tels que *' Ia si 'ia*, Asia, Asie ; *Ing ki li*, Englisch, Anglais ; *Fei li pe eul to*, Philibert. On sait également que c'est des caractères chinois envisagés au point de vue purement phonétique que procède l'écriture des Japonais, dont la langue est si différente de la langue chinoise.

Quant aux signes chinois eux-mêmes, nous avons déjà dit qu'ils avaient pour origine un véritable système d'imagerie. On les rencontre encore avec cette forme primitive sur certains monuments et on peut suivre leurs transformations graduelles à travers le cours des âges. Plusieurs systèmes graphiques ont été fixés d'une façon très précise, ont été employés durant des périodes de plusieurs siècles et n'ont dû qu'à des circonstances particulières de se voir plus ou moins sérieusement modifier.

D'ailleurs, il existe aujourd'hui encore chez les Chinois plusieurs sortes d'écritures, et parmi elles une espèce de

cursive assez rapide, qui est usitée dans les relations habituelles.

Nous ne nous étendrons pas davantage sur cette question des caractères chinois. C'est pour nous un sujet accessoire ; en effet, ce n'est point de graphique que nous nous occupons, c'est seulement de la structure et du matériel phonétique des langues.

§ 2. L'annamite.

L'annamite est la langue de l'Indo-Chine orientale. Au nord elle s'étend sur le Tonkin, au sud sur la Cochinchine.

L'annamite est séparé du siamois (au moins au sus-ouest) par un idiome dont le caractère n'est pas encore déterminé, le cambodgien. Nous engageons le lecteur à consulter la carte ethnographique de la partie sud-orientale de l'Indo-Chine dressée par Francis-Garnier (1).

La langue annamite est absolument distincte du chinois, et par son appareil phonétique et par ses racines, c'est-à-dire par ses mots, puisque la racine constitue le mot lui-même dans toute langue monosyllabique.

Tout comme en chinois, le genre et le nombre s'indiquent par l'adjonction, à la racine principale, de racines au sens de *mâle*, *féminin*, ou de *tous*, *nombreux*. L'adjectif se reconnaît à sa position après le substantif qu'il qualifie. La notion de temps ou de mode s'exprime enfin par l'emploi simultané de la racine sur laquelle pivote la phrase et d'autres racines dont le sens général est celui du passé, du futur, et ainsi de suite.

Ce que nous avons dit de la structure du chinois s'applique donc de point en point à l'annamite. Chez ce der-

(1) *Journal asiatique*, août-septembre 1872.

nier également, le système des intonations joue un rôle capital ; il distingue, comme en chinois, des mots dont la prononciation serait absolument la même, bien que leur sens soit tout à fait différent. Les intonations annamites sont au nombre de six : ton aigu, fort difficile à décrire ; ton interrogatif ; ton ascendant ou remontant, assez peu différent du ton interrogatif ; ton descendant ; ton grave ; ton égal.

L'écriture annamite est figurative, c'est-à-dire idéographique, et a été empruntée anciennement aux Chinois ; elle a subi, d'ailleurs, des modifications sensibles, et par la suite des temps l'on y a joint de nouveaux signes.

La langue annamite, au surplus, a fait au vocabulaire chinois des emprunts considérables, notamment au dialecte méridional ; ce fait a induit en erreur certains auteurs qui ont voulu comparer les deux idiomes et leur donner une origine commune. Le nombre, quel qu'il soit, de ces mots d'emprunt n'a rien à faire avec le fond même de la langue, avec ses racines propres ; celles-ci, fussent-elles même beaucoup moins nombreuses encore, suffiraient à établir l'originalité incontestable et l'indépendance de la langue annamite.

§ 3. Le siamois.

Le siamois occupe la région située au nord du golfe de Siam, assez avant dans l'intérieur du pays, et la côte occidentale de ce golfe. A l'est, il confine au cambodgien, idiome bien peu connu ; à l'ouest, il confine au birman, qui, lui aussi, est une langue monosyllabique.

On peut consulter sur les limites du siamois la carte annexée par M. Cust à son ouvrage *A Sketch of the modern Languages of the East Indies.*

Le nom de siamois, ou thaï (ou taï), est propre à une certaine population, dont le nombre d'individus ne dépasse guère le chiffre de 2 millions ou 2 millions et demi. La ville principale de ce domaine est Bangkok. Mais à côté du siamois proprement dit, on peut placer certains autres idiomes formant avec lui une famille linguistique.

C'est d'abord la langue des *Laos*, sujets de Siam, habitant les rives du fleuve Mékong (au nord-est du siamois proprement dit), et dont le nombre est évalué approximativement à 1 million.

Plus au nord, la langue des *Shans*, qui a reçu de l'ouest l'influence du birman.

Le *khamti* est parlé plus au nord encore, tout à l'extrémité nord-orientale du territoire de langue hindoue (assami) et isolé de toutes parts d'avec les autres idiomes de la famille siamoise. Le territoire d'Assam a été envahi au treizième siècle par des conquérants siamois, mais ceux-ci y perdirent bientôt leur langue et adoptèrent celle du pays qu'ils occupaient, l'assami.

On rattache également au siamois la langue des *Miao-tsé*, qui habitent dans quelques régions montagneuses de la Chine méridionale (1).

La phonétique du siamois est des plus riches ; on y rencontre notamment un grand nombre de consonnes aspirées et de sifflantes. L'alphabet est d'origine hindoue.

Quant à la structure même de la langue, elle est franchement monosyllabique ; la racine même constitue le mot. Les pronoms ne sont que des racines comme les autres, que l'usage affecte à leur rôle spécial. Il y a cinq tons différents, cinq façons d'accentuer les monosyllabes pour les diversifier. Comme en birman, comme en annamite, le système de numération est décimal.

(1) Edkins, *The Miau-tsi tribes*, Foochow, 1870.

Ajoutons que le lexique du style élevé ou religieux comprend un certain nombre de mots empruntés aux idiomes hindous.

§ 4. Le birman.

Il est parlé au nord-ouest de la péninsule indo-chinoise, entre le siamois et les langues hindoues. Son matériel phonétique est moins riche que celui du siamois ; on n'y compte qu'une sifflante. Les différents tons du birman semblent être moins nombreux que ceux du chinois et du siamois. Quant aux procédés grammaticaux, ils sont tout à fait les mêmes.

On trouve toutefois, en birman, une tendance à la dérivation au moyen de la préfixation. Des noms sont dérivés, en effet, au moyen d'un *a* placé avant la racine de *kaunh*, ayant le sens général de « bon, être bon », ou forme *a-kaunh* « bonté », etc. Pour tout le reste, la langue est franchement monosyllabique. Il est intéressant, en tout cas, de constater cette tendance évolutive.

En ce qui concerne les limites des différentes langues de l'Indo-Chine, nous renvoyons au volume de R. N. Cust, *A Sketch of the mod. Languages of the East Indies.*

§ 5. Le tibétain.

Le Tibet doit à l'Inde bouddhiste la meilleure part de sa culture intellectuelle, son alphabet, son importance littéraire. Il est difficile de savoir ce que pouvait être la littérature tibétaine avant le mouvement religieux qui sans doute la transforma entièrement. Nous n'avons point de documents remontant à cette première époque.

Les missionnaires bouddhistes eurent pour premier soin de traduire en tibétain les livres religieux composés

en sanskrit. L'alphabet qu'ils employèrent, et qui se trouve encore usité, était celui qui avait cours dans l'Inde septentrionale ; son origine est parfaitement évidente, et quiconque lit le caractère hindou dévanâgarî apprend en quelques heures l'alphabet tibétain, qui en provient directement.

Les différents auteurs qui ont écrit sur le tibétain n'ont pas mis suffisamment en lumière le caractère monosyllabique de cette langue. Les procédés qu'elle emploie sont analogues à ceux dont se servent le chinois, l'annamite et les autres langues isolantes.

C'est ainsi que le tibétain ne connaît dans les noms ni genre ni nombre. Pour exprimer le genre d'un nom, il doit le faire accompagner d'un autre mot dont le sens est celui de mâle ou de femelle : *ra pho* « bouc », *ra ma* « chèvre ». De même il ne peut exprimer le pluriel qu'en adjoignant au nom qui doit comporter cette idée de pluralité un autre mot dont le sens est, pour l'ordinaire, celui de *tout* ou de *multitude*.

Les prétendus cas du tibétain sont aussi peu des cas que ne le sont ceux qu'on attribue au chinois ou à l'annamite ; ici également on emploie pour déterminer la racine *pleine*, le mot *plein*, des mots qui deviennent *vides*, c'est-à-dire qui perdent une partie de leur sens premier et servent en quelque sorte d'adjoints au mot principal.

Par lui-même, le mot n'est pas plus un simple nom (substantif ou adjectif) qu'il n'est un verbe. C'est la position dans la phrase ou l'adjonction de telle ou telle racine dite *vide*, qui peut résoudre ce problème.

On ne peut nier, toutefois, que le tibétain ne montre de fortes tendances à l'agglutination, par le procédé de suffixation de racines secondaires aux racines principales. C'est, en somme, une langue que l'on saisit en état d'évolution.

Le système vocalique est assez simple, celui des consonnes assez riche.

La numération est décimale.

Après tout ce que nous avons dit des langues monosyllabiques en général, et du chinois en particulier, il nous semble inutile d'examiner d'une façon plus minutieuse la structure du tibétain. Elle n'est pas différente de la structure des autres langues isolantes, et il ne faut point se laisser prendre à ce que les grammaires ordinaires disent de ses prétendus genres, nombres, cas, personnes, temps et modes. Ce sont là tout autant de façons de parler qui ne doivent pas êtres prises à la lettre, et dont il n'y aura plus trace dans la syntaxe comparée des différentes langues monosyllabiques, qu'un avenir prochain verra sans doute paraître. Celui qui entreprendrait de réaliser cette tâche et la mènerait à bonne fin, sans tenter de réduire à une forme commune les racines tout à fait diverses de ces langues, aurait rempli l'un des premiers *desiderata* de la linguistique.

Mais, avant tout, il faudrait voir plus répandue cette idée que pour étudier une langue monosyllabique quelconque, il est nécessaire d'oublier momentanément ce que l'on sait de la structure et du mode de fonction de nos langues à flexion. Il paraît, malheureusement, que ce n'est point une petite difficulté.

Les régions situées au nord et à l'est de l'Inde sont le domaine d'un grand nombre d'idiomes, dont une bonne partie doivent être rangés, par leur structure, dans l'ordre des langues monosyllabiques, tout en donnant des témoignages non équivoques d'une tendance fort marquée à l'emploi des procédés d'agglutination. D'autres, classés à côté des précédents, devraient plutôt être regardés comme véritablement entrés dans la période agglutinante.

Quoi qu'il en soit, citons, immédiatement au nord du

Népaul britannique et dans la région himalayenne, immédiatement au sud du Tibet : le *gouroung*, le *magar*, le *mourmi*, le *néwar*, le *kiranti ;* puis, en continuant vers l'est, le *léptcha*, sur le territoire des Sikhs indépendants ; plus à l'est encore, le *kachari*, l'*aka*, le *dophlia*, le *miri*, l'*abor*, longeant tous cinq la frontière septentrionale de l'assami (idiome aryen) ; à la limite orientale de ce dernier (et tout contre le domaine du khamti, voir ci-dessus), le *michmi ;* dans la bande de territoire de langues non aryennes, rentrant entre l'assami et le bengali : le *garo*, le *mikir*. A l'est du bengali (et au nord du siamois), le *munipuri* et vingt ou trente autres idiomes. On peut consulter sur la géographie linguistique de toute cette région le livre déjà cité de R. Cust, excellent ouvrage de linguistique et d'ethnographie.

Le *khasia* est parlé entre l'assami (au nord), le bengali (au sud), le garo (à l'ouest), dans le pays montagneux qui sépare la vallée du Brahmapoutra (au nord) du Bengale oriental (au sud); il comprend jusqu'à six variétés dialectales, plus ou moins différentes les unes des autres, mais parmi lesquelles il est difficile de reconnaître la variété qui peut être regardée comme type. C'est la langue de deux cent mille individus environ, assez civilisés si on les compare à nombre d'autres populations du sud du Tibet et du nord de la Birmanie.

Cet idiome, toutefois, n'a aucune littérature propre, aucun système graphique. Les documents que l'on a recueillis, soit des récits populaires, soit des traductions des livres chrétiens, sont transcrits en caractères bengalis ou en caractères européens.

Le khasia est pour le linguiste d'un intérêt tout spécial, en ce qui concerne la structure. On peut trouver, sans doute, dans les contrées du Tibet méridional et de l'extrême Orient, plus d'un idiome présentant, comme le khasia, la phase de développement morphologique qui

répond au passage du monosyllabisme à l'agglutination; mais on n'en connaît pas, jusqu'ici, qui présente d'une façon plus frappante ce phénomène capital d'évolution. Nous n'insistons pas pour l'instant ; un peu plus loin, en effet, nous prendrons le khasia pour exemple lorsqu'il s'agira d'expliquer le passage de la phase linguistique, isolante, ou monosyllabique, à la phase agglutinante.

Au fond du golfe de Martaban, c'est-à-dire à l'embouchure de l'Irawaddi, du Sitang et du Salwen, est parlé le *mon* ou pégouan, ainsi nommé de la province dont il occupe une grande partie. Les villes de Pégu, Martaban, Moulmein Yey sont situées dans le domaine du mon ; le pays est habité par les Talains, de religion bouddhique. Cette langue est tout à fait distincte du birman qui l'avoisine et qui, vraisemblablement, l'a jadis repoussée au sud, dans la région qu'elle occupe actuellement.

Si nous remontons plus au nord, par delà même la ville d'Amarapoura, et toujours entre l'Irawaddi, à l'ouest, et le Salwen, à l'est, nous rencontrons le territoire assez peu considérable du *paloung*, confinant, à l'occident, au birman, et, à l'orient, à des dialectes siamois. La ville de Bamo est à une petite distance, au nord du territoire du paloung.

On regarde également comme appartenant par leur structure à l'ordre des idiomes monosyllabiques les différents dialectes parlés par les populations plus ou moins sauvages qui habitent au nord du Cambodge, entre le pays de Siam, à l'ouest, et l'Annam, à l'est, c'est-à-dire dans la région moyenne du fleuve Mékong : ce sont les Pnoms des Cambodgiens, les Moïs des Annamites, les Khous des Siamois. En tout cas, nous sommes ici sur un terrain encore bien peu connu.

Que dire aussi du *cambodgien*, parlé dans le territoire du Mékong mériodional par un million et demi d'indi-

vidus, regardé comme un dialecte siamois (idiome auquel il a d'ailleurs emprunté beaucoup de mots) ? En somme, les renseignements que l'on possède sont insuffisants. En tout cas, le cambodgien est parlé « recto tono » et ignore les tons qui jouent un si grand rôle dans les idiomes monosyllabiques.

CHAPITRE IV.

SECONDE FORME LINGUISTIQUE
L'AGGLUTINATION
LES LANGUES AGGLUTINANTES

Nous devons avant tout définir l'agglutination et rechercher quelle est son origine ; nous passerons ensuite en revue les principaux systèmes linguistiques qui revêtent cette forme particulière.

§ 1. Qu'est-ce que l'agglutination ?

Tandis que dans les idiomes de la première forme (le chinois, le siamois et les langues analogues), les mots ne sont autre chose que des formes monosyllabiques invariables, placées à la suite les unes des autres (sans qu'il y ait cependant entre elles une juxtaposition très intime), il arrive dans les idiomes du second degré que plusieurs éléments se juxtaposent réellement, s'agglutinent, s'agglomèrent ; de là, le nom de *langues agglutinantes* ou *agglomérantes* qui leur a été donné.

Les divers éléments qui entrent dans la confection du mot ne possèdent plus chacun leur valeur propre, leur valeur première. Il n'y en a plus qu'un seul qui porte l'idée principale, l'idée de la signification, le sens. Les autres éléments perdent tout à fait leur valeur indépendante. A la vérité, ils possèdent bien encore une portée personnelle, individuelle ; mais ce n'est qu'une portée toute relative. En effet, tandis que l'élément dont la signi-

fication aura persisté avec sa valeur primitive — « frapper, prendre, garder », et ainsi de suite — verra se grouper autour de lui des éléments qui détermineront les modes d'être ou les modes d'action, d'autres éléments, perdant de leur valeur primitive, s'accoleront à cet élément dont la signification est tout entière sauvée et auront pour rôle de déterminer les modes d'être ou d'action de l'élément en question.

Si nous représentons par R — lettre initiale du mot « racine » — l'élément ainsi sauvegardé, l'élément dont le sens a persisté tout entier, et si nous représentons par une série de lettres R les éléments qui sont tombés à la condition de simples éléments de relation, nous pouvons supposer dans une langue agglutinante les formes suivantes de mots : R R, soit la racine de signification précédée d'une préfixe, signe de relation — R R, soit la racine suivie d'un suffixe — R R R, soit la racine entre deux éléments de relation — R R R R, et ainsi de suite.

Placé avant la racine principale, l'élément de relation est appelé *préfixe :* c'est R dans la forme R R. Placé après cette même racine, il reçoit le nom de *suffixe ;* dans la forme RRRR nous trouvons un préfixe et deux suffixes.

Préfixes et suffixes reçoivent le nom général d'*affixes.*

Deux ou trois exemples rendront d'ailleurs plus saisissante l'explication que nous venons de donner. Nous les empruntons à la langue magyare.

Dans la forme *kértëk* « vous priez », kér est la racine, l'élément dont la signification entière est sauvegardée ; *tëk* est l'élément de rapport et indique la personne ; d'après ce qui vient d'être dit, la formule du mot est donc RR. S'agit-il de *kérnétëk* « puissiez-vous prier », nous avons la formule RRR : en effet, l'élément juxtaposé *né* n'est qu'un signe de relation indiquant que l'idée générale et dominante de *kér* « prie » est ici à l'optatif. Prenons la racine *zár* « fermer » et jetons les yeux sur

quelques-unes de ses formes soi-disant dérivées qui, en définitive, ne sont que des exemples d'agglutination, de juxtaposition. Elles nous laissent voir de la façon la moins équivoque ce que c'est en réalité que le phénomène dont nous nous occupons : *zárhat* « il peut fermer », formule RR ; *zárogat*, « il ferme souvent », même formule ; *zárogathat* « il peut fermer souvent », formule RRR ; *zárat* « il fait fermer », formule RR ; *záratgat* « il fait fermer souvent » formule RRR ; *záratgathat* « il peut faire fermer souvent », formule RRRR.

Deux faits caractéristiques distinguent donc la classe agglutinante de la classe monosyllabique. Dans la classe agglutinante, le mot n'est plus composé de la racine seule, mais il est formé de l'union de plusieurs racines. En second lieu, dans cette juxtaposition une seule des racines agglomérées garde sa valeur réelle : les autres racines voient leur signification individuelle s'amoindrir, passer au second rang ; elles ne servent plus qu'à préciser le mode d'être ou d'action de la racine principale dont la signification primitive est sauvegardée.

Nous avons dit un peu plus haut, en parlant des langues monosyllabiques, qu'un certain nombre des idiomes classés dans ce groupe étaient déjà entrés, en partie, sinon totalement, dans la phase de l'agglutination. Le cas est bien curieux pour la langue khasia. Le khasia, par sa structure, ne peut être rangé ni parmi les idiomes isolants appartenant à la première phase morphologique du langage, ni parmi les idiomes agglutinants appartenant à la seconde phase. Il offre un exemple très frappant de l'évolution linguistique, et nous permet de prendre sur le fait la transformation d'une langue isolante en langue agglutinante. Par de nombreux côtés, le khasia appartient au monosyllabisme. Ainsi, comme dans toute langue monosyllabique, le mot, quel que soit son genre, reste invariable, et il reste invariable également s'il cor-

respond à une idée d'unité (un homme) ou à une idée de pluralité (des hommes). Voici, par exemple, le mot *ksew :* s'agit-il d'un chien mâle, s'agit-il d'un chien femelle, s'agit-il d'un ou plusieurs chiens ? Il faut avoir recours à un procédé accessoire, à un artifice. On fait précéder le nom dont il faut déterminer le genre soit du mot *u*, qui signifie « il », soit du mot *ka*, qui signifie « elle ». Exemples : *u briw*, l'homme ; *ka briw*, la femme; — *u ksew*, le chien ; *ka ksew*, la chienne ; — *u kun*, le fils ; *ka kun*, la fille. Mais, sous d'autres rapports, le *khasia* appartient, en partie au moins à la seconde phase linguistique. Un certain nombre de racines *vides* (voir ci-dessus), au lieu de rester indépendantes, tout en jouant le rôle d'éléments de relations, s'agglutinent à la racine *pleine*, dont elles expriment les rapports, et deviennent ainsi de véritables éléments dérivatifs. Répétons-le, le mot dérivé se compose de deux (ou plusieurs) racines : l'une, principale, ayant conservé toute sa valeur radicale, sa valeur « pleine » ; l'autre, l'élément dérivatif, n'ayant plus qu'une valeur d'auxiliaire. Ainsi, dans les mots latins *pater*, *mater*, *frater*, la première syllabe est le radical et porte le sens principal ; la seconde, *ter*, ne joue qu'un rôle accessoire : c'est une ancienne racine « pleine » qui est devenue « vide » et s'est intimement soudée, pour mieux la dériver, à la racine qui s'est maintenue « pleine ». Lorsque l'élément dérivatif est après la racine qu'il s'agit de dériver, cet élément reçoit le nom de *suffixe ;* il y a dérivation par suffixe. Lorsqu'il précède l'élément radical (si, par exemple, l'on disait *terpa*, *terma*, *terfra*), il est dit *préfixe*, et il y a dérivation par préfixe. En khasia, nous trouvons la dérivation par préfixes et celle par suffixes, la première étant ici la plus fréquente. C'est le cas pour *nong*, qui forme nombre de mots appelés « noms d'agents » : *âp*, veiller ; *nongâp* veilleur ; — *bûd*, suivre : *nongbûd*, suivant ; — *ban*, pres-

ser ; *nongban*, oppresseur ; — *dih*, boire ; *nongdih*, buveur ; — *kam*, travailler ; *nongkam*, travailleur ; — *tuh*, voler ; *nongtuh*, voleur. Le mot *jing*, chose, a parfois aussi sa valeur pleine et entière. C'est un préfixe souvent employé : *bam*, manger ; *jingbam*, aliment ; — *duh*, détruire ; *jingduh*, dissolution ; — *thaw*, créer ; *jingthaw*, créature ; — *mut*, penser ; *jingmut*, pensée. Au moyen de l'élément *ba*, préfixé, on forme les mots ayant la valeur d'épithètes : *lih*, *balih*, blanc ; *iong*, *baiong*, noir ; *bha*, *babha*, bon. La dérivation par suffixes est moins importante en khasia, avons-nous dit, que celle par préfixes. Elle existe pourtant, témoins les dérivés par *ba*, tels que : *uba*, lequel; *kaba*, laquelle ; *kiba*, lesquels, etc. En somme, nous prenons ici sur le fait le passage de l'isolement, du monosyllabisme, à l'agglutination : le khasia (et c'est le cas de plusieurs autres idiomes asiatiques que l'on range parmi les langues agglutinantes) est en voie d'évolution, et démontre à l'évidence que l'espèce n'est qu'un moyen plus ou moins commode, de classification.

La racine principale maintenue dans sa forme primitive, les racines accessoires (si nous pouvons employer ce terme) perdant leur indépendance et se juxtaposant à la racine principale, voilà ce qui constitue l'agglutination. Le mot, ici, est formé par la réunion de plusieurs éléments divers, par la réunion de plusieurs racines ; il est complexe. C'est ce qui le distingue du mot tel que le conçoivent les langues isolantes, où il est formé de la racine elle-même, d'une seule racine.

Nous devons en tous cas le faire observer dès maintenant ; il n'y a pas encore dans les langues agglutinantes de vraie déclinaison, de vraie conjugaison. Si l'on se sert de ces mots de déclinaison, et de conjugaison, ainsi que des mots de cas, de nominatif, d'accusatif, de génitif, et ainsi de suite, en parlant du japonais, du basque, du

wolof, cela n'est qu'une façon de dire. Nous ne la blâmons pas absolument, mais nous tenons à établir nos réserves.

De toutes les langues connues, celles qui par leur forme appartiennent à la seconde classe, ou, pour mieux dire, à la seconde couche, sont de beaucoup les plus nombreuses.

Ce n'est pas à dire qu'elles soient apparentées les unes aux autres. Les étymologistes de profession, ceux principalement qui se rattachent à l'orthodoxie judéo-chrétienne, ont tenté mainte et mainte fois de les ramener à une origine commune, de leur trouver un fond commun. Leurs efforts n'ont obtenu que le succès malheureux dont ils étaient dignes. Certes, toutes les étymologies faites sans méthode permettront de rapprocher le magyar et le basque, le tamoul et l'algonquin, le japonais et les idiomes australiens. Mais qu'est-ce que l'étymologie en dehors de la grammaire ? un amas de fictions et de conceptions chimériques, un jeu d'esprit, un défi perpétuel aux principes les plus rudimentaires de la méthode et, le plus souvent, aux premiers éléments du bon sens.

Nous avons dit que le nombre des langues agglutinantes était considérable et qu'elles formaient la grande majorité des idiomes connus. Nous allons jeter un coup d'œil sur ces différents idiomes, ou du moins sur ceux d'entre eux qui paraissent représenter de la façon la plus frappante les principaux systèmes agglutinatifs.

Nous devrons passer rapidement sur quelques-unes de ces langues, le coréen, par exemple, ou certaines langues des nègres d'Afrique ; mais nous nous occuperons avec plus de détails de quelques autres idiomes appartenant à la seconde couche linguistique, tels que le basque, les langues dravidiennes, les langues américaines. On conçoit aisément pour quel motif nous nous arrêterons davantage sur celles-ci et moins sur celles-là.

Après avoir énuméré les principaux groupes de langues agglutinantes nous dirons quelques mots du « touranisme », des prétendues « langues touraniennes » et des conceptions imaginaires auxquelles cette théorie a donné naissance.

L'ordre dans lequel nous allons parler des langues agglutinantes est un peu arbitraire, au moins en ce qui concerne quelques-unes d'entre elles.

Nous commencerons par les langues agglomérantes de l'Afrique : langues des Hottentots, des Bochimans, des Nègres africains, des Cafres, des Pouls, des Nubiens. Poussant vers l'est, nous arriverons aux Négritos, aux Papous, aux Australiens. Remontant au nord-ouest, nous rencontrerons le système maléo-polynésien ; plus au nord, à l'extrême orient, le japonais et le coréen. Revenant vers l'ouest nous trouverons les langues dravidiennes dans le sud de l'Inde, le groupe ouralo-altaïque en Asie et en Europe, le basque au pied des Pyrénées occidentales, et, en traversant l'Atlantique, les langues américaines. Nous terminerons par les idiomes du Caucase et certaines autres langues peu connues ou dont la place n'est point déterminée.

La première partie de cette énumération est purement géographique ; mais nous avons suivi une certaine raison grammaticale en rangeant à la suite les uns des autres les idiomes dravidiens, ouralo-altaïques, basque et américains. Il nous serait difficile d'expliquer dès maintenant la cause qui nous a fait suivre cet ordre ; nous la ferons connaître en temps et lieu, notamment lorsque nous en arriverons à traiter des langues américaines.

§ 2. Langues de l'Afrique méridionale.

Nous faisons abstraction ici des idiomes appartenant au système « bantou », dont nous parlerons un peu plus

loin sous le nom de « langues des Cafres ». Par ce terme de langues de l'Afrique méridionale nous entendons les idiomes des Hottentots et ceux des Bochimans.

Langue des Hottentots.

La question de l'origine des Hottentots est tout à fait obscure. L'origine de leur langue n'est pas mieux connue. On a cherché à la rattacher aux langues khamitiques, à l'ancien égyptien, au copte ; cette tentative n'a pas eu de succès. Telle quelle, la langue des Hottentots nous semble isolée, indépendante de tout autre idiome.

Elle est d'ailleurs franchement agglutinante.

Les dialectes hottentots sont au nombre de trois : le *nama*, le *kora*, le *hottentot du Cap*.

Le premier de ces dialectes, le nama, est le plus important, des trois ; il serait parlé par vingt mille individus environ. Confinant vers le nord au domaine du héréro (langue du système bantou dont nous parlerons tout à l'heure), borné au sud par le fleuve Orange, le territoire du nama a pour limite occidentale l'Atlantique et pour limite orientale le désert de Kalahari (1).

Le kora est parlé beaucoup plus à l'est, dans l'intérieur des terres, dans la région des rivières Vaal, Modder et Caledon, aux environs du 29e degré de latitude. Ce dialecte, assez rapproché du précédent, est en voie d'extinction rapide.

Le hottentot du Cap est à peu près éteint. Il s'étendait sur le territoire de la colonie, confinant vers le nord-est à des idiomes du système bantou, vers le nord au kora, vers le nord-ouest au nama. Il n'y a plus aujourd'hui

(1) Th. Hahn, *Die Sprache der Nama*. Leipzig, 1870. — Tyndall, *A Grammar and Vocabulary of the Namaqua-Hottentot Language*. — Bleek, *A Comparative Grammar of the South Africa Languages*.

qu'un bien petit nombre de Griqouas qui se servent entre eux du hottentot ; le hollandais, le cafir l'ont presque entièrement étouffé.

Les différences d'ailleurs ne sont point considérables entre ces trois dialectes, et les Griqouas peuvent comprendre aisément le nama des bords de l'Atlantique.

Nous nous servons des mots de nama et de kora avec intention à l'exclusion de ceux de namaqoua et de korana. Ces deux derniers en effet ne sont que des mots dérivés.

Le Hottentot, dans sa propre langue, s'appelle Khoïkhoïb, au pluriel Khoïkhoïn ; ce nom a le sens de « homme des hommes » ou « ami des amis » (Hahn, *op. cit.*, p. 8).

La phonétique du nama est fort variée ; la série de ses voyelles est très-nuancée et il peut les nasaliser toutes. Il possède également un assez grand nombre de diphthongues, une douzaine, dit-on.

Il n'est pas moins riche en consonnes. Outre les explosives ordinaires (*p*, *t*, *k* et *b*, *d*, *g*), il possède plusieurs gutturales (*kh* et autres) ; les sifflantes *s* et *z* (de « sœur » et de « zèle ») ; une nasale particulière qui équivaut à peu près à la nasale de l'allemand « enge » ; *v*, *r*, *h*, et une palatale (qui ne se rencontre pas, d'ailleurs, dans le dialecte nama).

A ces consonnes diverses nous devons en ajouter quatre autres d'un ordre tout particulier, les « clics » ou claquements, les consonnes claquantes. Le claquement *dental*, figuré par un trait vertical : ı (ou, selon quelques auteurs, par un *c*) ; le claquement palatal, figuré par deux traits horizontaux coupant un trait vertical (ou par la lettre *v* d'après quelques auteurs ; le claquement *cérébral*, rendu par un point d'exclamation : ! (ou par la lettre *q*) ; le claquement *latéral* figuré par deux traits verticaux : ıı (ou par un *x*). Ces consonnes particulières

sont bizarres pour une oreille européenne, mais on arrive pourtant à les imiter. On trouvera leur description dans les grammaires spéciales, du moins celles des trois premiers. Le quatrième est fort étrange et reçoit son nom de ce que les dents latérales jouent un rôle important dans son articulation.

Les claquements peuvent précéder les consonnes gutturales, *n*, *h*, et toutes les voyelles ; on les rencontre d'ailleurs à chaque instant, presque à chaque mot.

Arrivons à la formation des mots qui est des plus simples : la racine suivie d'un suffixe, c'est-à-dire d'un élément dérivatif.

Tout d'abord notons que ces éléments dérivatifs, que ces suffixes ont chacun une triple forme ; l'une est réservée pour le cas où le mot est sujet, l'autre pour le cas où le mot est régime (direct ou indirect, peu importe). La première forme reçoit le nom de *subjective*, la seconde celui d'*objective*. La troisième forme est celle du *vocatif*, c'est la forme dite *interjective*.

Notons ensuite que les suffixes ont une forme pour le singulier, une forme pour le duel, une forme pour le pluriel ; cela fait donc neuf formes pour un seul et même élément, puisqu'il peut être subjectif singulier, objectif duel, interjectif pluriel et ainsi de suite.

D'autre part, nous nous trouvons en présence d'une triple hypothèse : l'élément dérivatif de la racine peut être un élément de la première personne (moi, nous deux, nous), ou un élément de la seconde personne (toi, vous deux, vous), ou bien encore un élément commun, un élément de troisième personne. Ce fait détermine quel est le suffixe du mot.

Dans les deux premières hypothèses, on forme des mots ayant, par exemple, le sens de « moi roi, moi qui suis roi », « toi qui est reine » et ainsi de suite. Répétons-le, d'ailleurs, l'élément change neuf fois pour un seul et

même mot, alors qu'il s'agit de la forme subjective, de la forme objective, de la forme vocative et selon qu'il est question du singulier, du duel, du pluriel.

Ajoutons encore que le suffixe varie selon que l'individu est du genre masculin, du genre féminin ou du genre neutre.

Des mots dérivés par un élément indiquant la première ou la seconde personne, passons à ceux que dérive un suffixe commun, un suffixe impersonnel. En nama (nous ne nous occupons que de ce dialecte) nous trouvons les désinences suivantes :

		Masculin.	Féminin.	Neutre.
Singulier	subjectif. . . .	*b*	*s*	*i*
	objectif	*ba*	*sa*	*ê*
Duel	subjectif. . . .	*kha*	*ra*	*kha* ou *va*.
	objectif	*khá*	*ra*	» »
Pluriel	subjectif. . . .	*gu*	*ti*	*n*
	objectif	*gâ*	*te*	*na*

En jetant un coup d'œil sur ce tableau, nous voyons tout d'abord que le mot *taras* « femme » est un féminin, singulier, subjectif. Ayant à rendre cette expression : « je vois la femme », nous nous servirons de *tarasa ;* dans cette autre phrase : « les deux femmes disent », nous nous servirons de *tarara*, et ainsi de suite. La forme *khoib* « homme » sera employée dans ces phrases : « l'homme dit, l'homme frappe » ; la forme *khoigu* dans celles-ci : « les hommes disent, les hommes frappent » ; la forme *khoigâ* dans celles-ci : « ils frappent les hommes, ils voient les hommes ». Tout ce mécanisme demande sans doute un peu d'attention, un peu d'habitude, mais il se laisse saisir assez facilement.

La dérivation secondaire s'opère d'ailleurs par l'adjonction de nouveaux suffixes aux suffixes qui dérivent déjà la racine, et c'est également à l'aide de nouveaux éléments annexés à la fin du mot que l'on exprime les

notions du locatif, de l'ablatif, de l'instrumental, et ainsi de suite. C'est par ce même procédé de dérivation que l'on forme des mots adjectifs tirés des mots substantifs.

Les formes causatives, diminutives, désidératives, intensives, sont tout autant de formes dérivées par l'adjonction à la racine principale de racines secondaires, c'est-à-dire d'éléments dérivatifs. Quant aux prétendues formes verbales, elles consistent simplement dans l'agglomération d'éléments dont l'un indique la personne (moi, toi, nous, ils), un autre la racine principale, le radical, un autre enfin, le temps (à présent, jadis, dans l'avenir).

Pour en finir avec le hottentot, disons qu'à la façon des langues monosyllabiques, il distingue ses mots homophones, en les chantant en quelque sorte sur des tons différents. Ces tons sont au nombre de trois : le mot *kaib*, par exemple, signifie « obscurité », ou bien « lieu », ou bien « linge », selon qu'il reçoit telle ou telle intonation. Ces homophones, d'ailleurs, ne sont pas très-nombreux.

D'autre part, ajoutons que l'accent, le véritable accent, tombe toujours en hottentot sur la syllabe radicale du mot, c'est-à-dire sur la première syllabe, puisqu'ici la forme du mot est celle-ci : racine + suffixe ou racine + suffixes. S'agit-il d'un mot composé, c'est-à-dire de deux mots réunis pour n'en faire qu'un seul, pour faire un mot complexe, l'accent appartient au mot principal.

Langues des Bochimans.

Les Bochimans, dispersés en un grand nombre de tribus peu considérables, ne donnent pas un nom général à l'ensemble de leur population. Les Hottentots les appellent *sân*, c'est-à-dire aborigènes, indigènes. Quant au nom de Bochimans, il est d'origine hollandaise et veut dire *homme des buissons*.

On ne connaît que fort peu de chose des divers idiomes parlés par les Bochimans. S'ils sont tous alliés les uns aux autres, il existe du moins de grandes différences entre tels et tels d'entre eux. On a voulu les assimiler aux dialectes hottentots, mais cette tentative n'a pas eu de succès ; tels que nous les connaissons, les idiomes des Bochimans sont indépendants de la langue des Hottentots. En tout cas, ils appartiennent comme elle au système agglutinatif. Ils connaissent plus de consonnes claquantes que les dialectes hottentots ; six ou sept, assure-t-on.

Le pays des Bochimans est assez difficile à délimiter. On les rencontre à l'est du territoire héréro, au nord-est du pays des Namas, au nord du désert de Kalahari. Au sud de ce même désert et de la rivière Orange, des Bochimans habitent le nord-ouest de la colonie du Cap. En somme, d'après Fritsch, ils se seraient étendus sur toute l'Afrique du sud, depuis le Cap jusqu'au Zambèse, et même par-delà ce fleuve (1). Ils auraient été chassés par la force des pays qu'ils n'occupent plus actuellement dans cette vaste région.

§ 3. Langue des Nègres d'Afrique.

L'Afrique est occupée au nord par un idiome sémitique, l'arabe, et un idiome khamitique, le berber. A l'est, en Abyssinie, on trouve également des langues sémitiques plus particulièrement alliées à l'arabe, et plus au sud, c'est-à-dire immédiatement au nord de l'équateur, quelques langues khamitiques classées sous le nom général de langues éthiopiennes. Tout le sud-est de l'Afrique et une forte partie de la côte sud-ouest sont occupés par les idiomes des Cafres, qui forment un groupe bien distinct.

(1) *Die eingeborenen süd-Afrika's*. Breslau, 1872.

Au sud se trouvent les langues des Bochimans et des Hottentots. Au centre même de la péninsule, de l'est à l'ouest, en partant du midi de la haute Egypte, on rencontre les dialectes nubiens et le poul, qui n'ont rien de commun avec les langues que nous venons d'énumérer.

Le reste de l'Afrique, c'est-à-dire la partie moyenne de la côte occidentale et une grande partie du centre, appartient aux idiomes parlés par les Nègres, par les véritables Nègres, que l'anthropologie ne confond pas avec les Cafres.

Le nombre des idiomes parlés par les Nègres d'Afrique est assez important. Quelques-uns de ces idiomes se rattachent d'assez près les uns aux autres et forment ensemble des groupes bien marqués ; mais on ne peut assurer, avec preuves scientifiques en mains, que ces différents groupes soient tous issus d'une seule et même souche. Ces différentes langues, sans doute, appartiennent les unes et les autres à la classe des langues agglutinantes, mais ceci ne préjuge en rien une communauté d'origine. Malgré bien des emprunts, le lexique de ces différents groupes d'idiomes est fort varié, et, par-dessus tout, leur grammaire est très diverse. Dans l'état actuel de nos connaissances, nous pouvons dire que l'on rencontre chez les Nègres d'Afrique un certain nombre de langues ou de groupes de langues tout à fait distincts les uns des autres, tout à fait indépendants.

Fr. Müller en compte 24 (*Allg. Ethn.*, 2e édit., 20). Ce chiffre est-il trop élevé, et des recherches ultérieures le feront-elles réduire ? N'est-ce là, au contraire, qu'un chiffre minimum, et découvrira-t-on quelque jour, parmi ces populations, des idiomes encore inconnus et qui ne rentreront point dans ces vingt-quatre familles ? C'est ce que nous ne pouvons prévoir. Contentons-nous d'insister sur ce fait, que cette expression de Langues des Nègres d'Afrique, qui forme le titre du présent paragraphe, est

purement géographique, et qu'elle n'éveille aucune idée de parenté entre les langues en question.

Nous procéderons à leur énumération, en nous conformant autant que possible à leur position géographique, du nord au sud et de l'ouest à l'est.

Le WOLOF. On possède une certaine quantité d'écrits sur la grammaire du wolof. Les formes de cette langue sont bien connues, et son lexique l'est suffisamment ; cependant, tous les travaux auxquels elle a donné lieu manquent de méthode et de critique. On a les éléments d'une grammaire wolofe scientifique, mais cette grammaire est encore à rédiger, et on ne peut guère la demander aux missionnaires qui habitent les contrées où cet idiome est parlé. Leurs nombreuses publications sont marquées au coin de la plus complète ignorance des procédés de la science moderne du langage, et ils ne paraissent point se douter de ce que c'est qu'une langue agglutinante.

Le système de voyelles du wolof est assez riche. A côté des voyelles brèves *a*, *é* (notre « é » fermé), *i*, *o*, *u* (notre voyelle « ou »), *è*, il possède des *â*, *î*, *ô*, *û*, *è* prolongés et un *é* fermé également long. Il connaît de plus un e qui paraît équivaloir à notre voyelle « e » de « que, je, te, le », et un *ă* bref et sourd qui, aux oreilles de ceux qui l'ont entendu, paraît intermédiaire entre notre « a » et notre « e » ; c'est vraisemblablement notre « e» prononcé d'une façon étranglée. Certains auteurs le rendent par « æ », mais ce procédé est manifestement défectueux. Dans un petit nombre de mots, le wolof possède un « a » nasal, correspondant à notre voyelle « an » de « grand, sang » ; mais, en général, la voyelle suivie de « n » se prononce sans nasalisation. Le wolof possède le son *ü* (notre « u » de « tu, lu »), mais ce n'est que dans des mots qu'il a empruntés au français. — Le wolof est également riche en consonnes. Outre les trois paires d'explo-

sives simples (*k* et *g*, *t* et *d*, *p* et *b*), il y a un « t » et un « d » mouillés, que nous transcrirons *t'* et *d'* ; les nasales *m*, *n*, *n'* (« n » mouillé, notre « gn »), et une nasale dite gutturale qui peut se trouver au commencement des mots, tout comme au milieu ou à la fin ; une aspirée très douce, *h*, et une gutturale *h°* répondant au « ch » allemand de « nach, noch »; *y* ; *r*, *l* ; la sifflante *s* dure et un *z* pour les mots empruntés au français ; la sifflante *f* et un *w* assez difficile à saisir pour nos oreilles européennes. Les groupes *mp*, *mb*, *nt*, *nd*, *ng* sont très fréquents, mais ce ne sont que des groupes de consonnes, non point des consonnes particulières.

Les mots correspondant à nos noms, soit substantifs, soit adjectifs, sont naturellement indéclinables, comme dans tous les idiomes appartenant à la classe de l'agglutination, et les désinences du latin, du grec et des autres langues à flexion, sont remplacées par des particules, par des prépositions. Cependant, lorsqu'il s'agit d'indiquer un régime direct et un régime indirect « donner un vêtement à Pierre », notre « à » ne s'exprime pas ; on a recours ici au procédé purement syntaxique, au procédé des langues isolantes, en un mot, à la façon de placer le mot dans la phrase ; ici on pose le régime indirect avant le régime direct. S'agit-il d'un nom qui est en état de dépendance vis-à-vis d'un autre nom (par exemple « roi, maître » dans ces propositions « le fils du roi, l'œil du maître »), ce nom est placé à la suite du nom principal, mais entre les deux est intercalé le conjonctif *u* que parfois, cependant, l'on sous-entend.

S'agit-il de désigner expressément le genre d'un nom, on lui adjoint un autre nom signifiant « mâle » ou « femelle », en rattachant ce qualificatif au mot qualifié par l'intermédiaire d'une particule exprimant la relation. — La forme du mot est d'ailleurs invariable et ne trahit en rien l'idée du singulier ou celle du pluriel. C'est une

particule, *i*, qui rend l'idée de ce dernier nombre. S'il est question de mettre au pluriel un nom ayant un complément, cette particule est intercalée entre les deux mots et remplace la particule *u* dont nous avons parlé ci-dessus et qui est réservée au singulier.

Le nom wolof est très souvent accompagné d'une particule qui lui est suffixée et qui joue le rôle d'un déterminatif. Cette particule est composée d'une consonne et d'une voyelle. La consonne varie d'après une loi euphonique, selon que le mot à déterminer commence par telle ou telle consonne : ainsi l'on dit *bay-bă* « le père », *făs-vă* « le cheval », *kăr-gă* « la maison ». Quant à la voyelle qui termine cette particule, elle varie, selon que le mot déterminé est présent (*i*), qu'il est proche, mais non présent (*u*), qu'il est éloigné (*ă*), qu'il est très éloigné (*â*). Ainsi le mot *kăr-gă*, que nous venons de citer, laisse entendre que l'on parle d'une maison déterminée, mais que cette maison est éloignée ; s'il s'agissait de la maison contre laquelle on se trouve, l'on dirait *kăr-gi*. S'il faut indiquer le pluriel d'un nom déterminé, la particule suffixée est, suivant les quatre hypothèses du plus ou moins de distance, *yi*, *yu*, etc. (en certains cas *n'i*, *n'u*, etc., avec « n » mouillée : *kăr-yi* « les maisons près desquelles on se trouve ». Cette particule *yi*, *yă*, *yu*, indice du pluriel, contient évidemment le signe pluriel *i*, dont nous avons parlé ci-dessus, et nous pouvons en conclure que les particules du singulier *gi*, *bă ku* et autres, le seul élément déterminatif est la voyelle ; mais quel rôle y joue la consonne initiale, c'est ce que nous ne savons pas encore.

Grâce à ce qui vient d'être dit dans ces quelques lignes, nous pouvons déjà nous rendre compte des propositions élémentaires telles que celles-ci : *făs u bûr* « cheval de roi ; *făs u bûrbă* « le cheval du roi » ; *făs u bûr-yă* « le cheval des rois » ; *făs i bûr* « chevaux du roi » ; *făs i bûr-bă* « les chevaux du roi » ; *făs i bûr-yă* « les che-

vaux des rois ». Si l'on fait abstraction de l'élément déterminatif de cette particule finale, on voit que ce procédé est très élémentaire et se saisit facilement. Le premier nom, comme l'on voit, ne prend pas le signe déterminatif ; naturellement, si le second nom n'est point déterminé, ni l'un ni l'autre ne le prennent : *făs u bûr, dah u nag* « beurre de vache ».

Il y a en wolof une façon de déterminer le mot de plus près encore, c'est de placer la particule déterminante avant ce mot, non plus après : *bi-bây, bă-bây, bu-bây* « ce père » ; ou encore de suffixer au mot déjà déterminé par le procédé habituel (*bây-bi, bây-bu*, etc.) la particule *lé : bây-bi-lé, bây-bu-lé*, etc. « ce père ». On peut même dire *bi-lé-bây, bu-lé-bay*, etc. Au pluriel on a, comme de juste, *yi-bây, bây-yi-lé*, etc. « ces pères ».

Il va de soi qu'à proprement parler, il n'y a pas plus de verbe, en wolof, qu'il n'y a de nom, en d'autres termes, que le mot ne s'y conjugue pas plus qu'il ne s'y décline. Les formes dites verbales que présentent dans leurs tableaux sans fin les grammaires wolofes, rédigées sur le modèle des grammaires latines et des grammaires grecques, ne consistent qu'en une agglomération de mots indépendants juxtaposés les uns aux autres. C'est le fait de toute langue agglutinante. La racine conserve toujours sa valeur tout à fait générale et des particules dont le but est d'exprimer l'idée du passé, celle du futur, celle du conditionnel, celle du subjonctif, etc., enfin les différentes idées des temps et des modes des langues à flexion, viennent s'adjoindre à cette racine, tantôt la précédant, tantôt la suivant. Rien ne varie dans cette agglomération, tous les mots juxtaposés restent les mêmes : il ne s'agit dans cette soi-disant conjugaison que de substituer les uns aux autres les pronoms « je, tu, il », etc. ; ces pronoms, d'ailleurs, se placent, selon les circonstances, en différents endroits de cette agglomération de mots.

Le nombre de ces combinaisons est considérable : les deux tiers de toute grammaire wolofe sont ordinairement consacrés à la prétendue conjugaison. En somme, il ne s'agit ici que d'apprendre à connaître la valeur d'un certain nombre de mots accessoires, de particules, et la place à laquelle on doit les poser dans l'agglomération qui constitue les mots. Ainsi la particule *on*, qui exprime l'idée de notre imparfait se place après le mot principal et avant le pronom personnel : *măs-nâ* « j'ai », *măs-on-nâ* « j'avais ». Mais ceci n'est qu'un exemple isolé, un exemple des plus simples ; à première vue les formes sont ordinairement très-compliquées et elles comprennent souvent six, sept, huit éléments et plus : *măs ăgu-nu-won-sopă-sopă-lu* « nous n'avions pas encore fait semblant d'aimer » n'est qu'une seule et même forme composée de différents mots agglutinés, de façon à n'en faire plus qu'un et ayant tous un rôle fixe, une position fixe dans cet assemblage. Les trois derniers éléments ont le sens de « ne pas faire semblant d'aimer » ; le premier, *măs*, indique l'action elle-même, *ăgu* dit qu'une action n'est pas encore commencée, *nu* est l'élément personnel, *won* est signe de l'imparfait. Ajoutons que ce mot n'est pas des plus compliqués ; nous pourrions en citer une foule d'autres qui semblent bien autrement touffus, mais le procédé de formation est toujours le même.

De toutes les langues des Nègres d'Afrique, le wolof est une des plus importantes, au point de vue des intérêts de la civilisation européenne. Les établissements français du Sénégal sont en contact journalier avec les Wolofs, et ceux-ci ont emprunté à notre langue un certain nombre de mots. Le long du fleuve du Sénégal, le wolof confine à la langue arabe parlée sur la rive droite de ce cours d'eau, et il s'étend au sud sur une grande partie de la Sénégambie : le wolof est la langue du Dyo-

lof, du Kayor, du Walo, du Dakar, et on le parle également dans le Baol, le Sine et la Gambie.

GROUPE MANDÉ. Le *mandingue* occupe la moitié méridionale de la Sénégambie et le territoire de la haute Guinée ; le *bambara* est parlé un peu plus au nord, à l'est de la Sénégambie centrale ; le *sousou*, le *véi*, le *téné*, le *gbandi*, le *landoro*, le *mendé*, le *gbésé*, le *toma*, le *mano* font partie de la même famille.

Le GROUPE FELOUP occupe également la Sénégambie méridionale et les régions situées un peu plus au sud ; il est en contact de divers côtés avec le mandingue dont nous avons parlé ci-dessus. Cette branche comprend de nombreux idiomes : le *feloup*, sur la Gambie, le *filham* sur le fleuve Casamanze, le *bola*, le *sérère*, le *pépel* dans les îles Bissagos, le *biafada* sur le fleuve Géba, le *padjadé*, le *baga*, le *kalloum*, le *temné*, le *boullom*, le *cherbro*, le *kissi*.

Le SONRAÏ est isolé. Il occupe la région du fleuve Niger, dans la partie de son cours située le plus au nord-est (au sud-est de Tombouctou), vers le 15e degré de latitude septentrionale. Le sonraï est donc parlé dans une partie du Sahara du sud, et son territoire confine à celui des Touaregs qui s'étendent plus au nord. On peut dire d'une façon générale qu'il est parlé de Tombouctou à Agadès.

Le HAOUSA, dont les dialectes sont nombreux, est en quelque sorte la langue du Soudan. Aucun autre idiome de l'Afrique centrale n'est aussi répandu que la haousa ; son territoire, au sud-est du souraï entre le Niger et le pays de Bornou, est fort étendu ; c'est la langue commerciale de l'Afrique du centre. Le haousa est assez bien connu, grâce notamment aux écrits du missionnaire anglais James F. Schön.

Les voyelles du haousa sont assez nombreuses. Outre les *a*, *i*, *u* (le son « ou » du français) et leurs correspondantes longues, un *o* et un *e*, il possède un *e* et un *i* exces-

sivement brefs (qu'il est assez difficile de distinguer l'un de l'autre) ; une labiale tenant le milieu de « a » et de « o » et qui peut être prolongée ; enfin un « a » et un « e » sourds et gutturaux. Cette échelle des sons est assez nuancée et donne au langage une certaine variété. Quant au système des consonnes, il n'a rien de compliqué. A côté des trois paires d'explosives ordinaires (*p*, *t*, *k*, et *b*, *d*, *g*), à côté des nasales *m*, *n*, des vibrantes *r*, *l*, des sifflantes *f*, *s*, *z*, *s* (le « ch » français, le « sh » anglais), *i* (le « j » français), des chuintantes que nous transcririons en français par « tch » et « dj », il possède une demi-voyelle *w* (dont le son paraît être celui de notre « u » dans nuit, suite »), et une nasale analogue à celle de l'anglais « king ».

Le genre, en haousa, peut être distingué, non-seulement par l'annexion au mot principal d'un mot accessoire dont le sens est celui de « mâle » ou de « femelle » (enfant+mâle=garçon, enfant+femelle = fille, et ainsi de suite), mais encore par une terminaison *ia* ou *nia* dont le sens n'est pas bien éclairci : *sa* « taureau »,*sania* « vache ». L'origine de cette terminaison doit évidemment être la même que celle de l'autre procédé. Le pluriel d'un nom est indiqué de même par l'annexion d'une particule (il y en a de plusieurs espèces) et parfois on redouble la dernière syllabe du mot. Dans la pratique cette formation du pluriel présente certaines difficultés, mais au point de vue de l'anatomie de la langue, elle n'a rien que de très simple et de très compréhensible.

Point de déclinaison véritable, point de cas, ainsi d'ailleurs que dans toutes les autres langues agglutinantes. C'est par sa position dans la phrase ou à l'aide de particules qui lui sont adjointes que le mot indéclinable prend la valeur des différents cas du grec et du latin : *ma-sa* « à lui », *ma-ta* « à elle », *gare-sa* « de lui, venant de lui ». Quant au mot qui est sujet de la phrase (dominus) et

quant à celui qui est régime direct (dominum), ils se trouvent désignés par leur place même : le dernier est naturellement posé après le premier. S'agit-il, enfin, d'exprimer la dépendance d'un mot par rapport à un autre mot (« le nom de pays, la sœur du père » et ainsi de suite, en un mot la notion du génitif grec et latin) le mot principal précède immédiatement l'autre mot, ou bien on place entre les deux mots la particule *na*, *n* au masculin, *ta* au féminin.

Comme dans toutes ses autres langues agglutinantes, c'est par l'accumulation de mots passés à l'état de particules, que se forment les prétendus temps et modes du haousa. C'est affaire aux grammaires spéciales que d'énumérer ces particules et d'expliquer leurs différentes manières de s'apposer les unes et les autres au mot principal. A première vue tout ce système semble un peu compliqué, mais il n'offre point de difficultés dont une analyse un peu méthodique n'ait aisément raison.

Le GROUPE BORNOU, ou bournou, est situé aux alentours du lac Tchad, dans l'Afrique centrale, à l'est du haousa dont nous venons de parler. Il comprend une demi-douzaine d'idiomes, parmi lesquels le *kanem* et le *téda*, langue des Tibbous, au nord et au nord-est du lac, le *kanori*, le *mourio*, le *ngourou*.

Le GROUPE KROU (*krou* et *grebo*) nous ramène sur la côte de l'Atlantique, près du fleuve Saint-Paul.

Le GROUPE EGBÉ, ou *évé*, occupe les régions situées vers la partie occidentale du golfe de Guinée, par le 7e degré de latitude et encore un peu plus au nord. On y compte quatre idiomes parents les uns des autres : l'*egbé*, le *yorouba*, l'*odji*, le *ga* ou *akra*.

L'IBO, autre rameau guinéen, est parlé dans le pays des embouchures du Niger. L'*ibo* est plus au sud, le *noupé* plus au nord.

Un peu plus à l'est, par le 7e degré de latitude, se trouve le MITCHI, idiome isolé.

Plus à l'est encore, au sud du groupe bornou et du lac Tchad, est situé le GROUPE MOSGOU : *mosgou, batta, logoné.*

Le BAGHIRMI, encore plus vers l'orient, au cœur même de l'Afrique, s'étend au sud-est du lac Tchad.

Le MABA est parlé plus avant encore dans la même direction et ne se rattache pas davantage aux idiomes qui l'environnent.

Enfin à l'est de l'Afrique centrale, placé au sud de la Nubie et à l'ouest de l'Abyssinie, se trouve un autre groupe de langues parlées également par les Nègres, le GROUPE des langues du HAUT NIL : le *chilouk*, sur la rive gauche du Bahr el Abiab ; le *dinka*, sur la rive droite du même fleuve ; le *nouer*, immédiatement au-dessous du chilouk ; le *bari*, vers le 5e degré de latitude et encore plus au nord.

Répétons-le avant de terminer, les différents groupes de langues parlées par les Nègres d'Afrique, par les Nègres de la Sénégambie, du Soudan et de la Guinée supérieure, sont généralement indépendants ; ils ne constituent pas une seule et même famille, bien qu'étant tous agglutinants.

En ce qui concerne la géographie de ces langues, nous renvoyons particulièrement au premier volume de l'ouvrage de R. N. Cust, *A Sketch of the modern Languages of Africa.*

§ 4. Langues du groupe bantou.

Le domaine de ces langues est considérable : on peut dire, d'une façon générale, qu'elles occupent le sud de

l'Afrique, abstraction faite des contrées où l'on rencontre les Bochimans et les Hottentots. Au sud elles atteignent les environs du Cap ; au nord, elles confinent au groupe éthiopien des langues khamitiques, aux langues des Nègres de Guinée et dépassent un peu la ligne équatoriale. Leur étendue en longueur correspond ainsi à la moitié totale de l'Afrique.

Un quart environ des Africains parlent les différents idiomes de cette famille. Les dialectes du groupe bantou sont nombreux et remontent tous à une origine commune. Nous avons vu qu'il était loin d'en être ainsi pour les langues parlées par les Nègres africains, au centre et à l'ouest de la péninsule. La langue mère qui a donné naissance aux différents idiomes de ce groupe est tout à fait inconnue, mais il n'est nullement impossible que l'on arrive un jour ou l'autre à en reconstituer les traits essentiels. Cette reconstitution portera aussi bien sur le lexique que sur le système grammatical.

Le nom général de langues des Cafres, que l'on donne parfois aux idiomes du groupe bantou, est un nom conventionnel. Ce mot de « cafre », qui est d'origine sémitique et veut dire « infidèle », après avoir été appliqué à toutes les populations du sud-est de l'Afrique, s'est trouvé limité de plus en plus. On ne le donne guère aujourd'hui qu'aux tribus qui s'étendent du nord-est de la colonie du Cap à la baie de Délagoa.

Il y a donc lieu de faire quelque réserve lorsqu'on donne le nom d'idiome cafre, soit au kisouahili dans le pays de Zanzibar, soit au dialecte de Fernando-Po dans le golfe de Guinée.

Le mot de « bantou » est préférable. C'est le pluriel du mot qui signifie « homme » ; il a le sens « d'hommes », de « population », de « peuple » et peut facilement s'appliquer par extension à la langue elle-même.

La phonétique de toute cette famille est des plus riches

et ne manque pas d'harmonie. En principe, les mots y sont formés par la préfixation — et non la suffixation — des éléments destinés à indiquer les relations et les modes d'être de la racine principale.

On divise en trois branches les langues du groupe bantou : une branche occidentale, une branche centrale, une branche orientale, et ces trois branches se divisent à leur tour en différents rameaux. Voici leur énumération sommaire d'après la classification généralement adoptée (Hahn, Fr. Müller) :

Branche de l'est : langues du pays de Zanzibar ; langues de la région du Zambèse ; groupe cafir-zoulou.

Branche centrale : sétchouana et tékéza.

Branche de l'ouest : congo, héréro, etc.

Les principales langues du groupe nord-est (région de Zanzibar) sont : le *kipokomo*, un peu au sud de l'équateur ; le *kisouahili* (par le 5e degré de latitude sud) ; le *kinika ;* le *kikamba ;* le *kihiaou*, vers le 13e degré. Le peuple le plus généralement connu d'entre ceux qui se servent de ces idiomes est celui des Souahilis.

Un peu plus au sud, nous trouvons les langues du Zambèse, *tété*, *séna* et autres.

Le *makoua*, un peu plus au nord-est, est parlé dans le pays de Mozambique.

Plus au sud encore, le *zoulou* et le *cafir*, fort rapprochés l'un de l'autre. (Consulter l'important ouvrage de R. N. Cust, *A Sketch of the modern Languages of Africa*, tome II, p. 1-434, avec la carte ; Fr. Müller, *Grundriss der Sprachwissenschaft*, tome 1er, p. 238 ss). Le premier de ces idiomes est parlé par les Amazoulous dans le pays zoulou et la terre du Natal ; le second par les Amakhosas ou Cafres proprement dits, au sud du territoire de Natal. Au cafir et au zoulou se rattache aussi le *fingou*, parlé par les Amafingous, les Amazouazis et quelques autres peuplades nombreuses. Ce groupe des Cafres s'étend

ainsi de la colonie du Cap jusqu'à la baie de Délagoa.

Des deux langues du groupe central, le *tékéza* est la moins connue.

L'autre, le *sétchouana*, l'est beaucoup mieux. C'est la langue des Bétchouanas, parlée plus au nord que le 20e degré de latitude, plus au sud que le 25e. Il comprend, à l'est, le *sésouto*, langue des Basoutos ; à l'ouest, le *sérolong*, le *sétlapi*, langues des Barolongs, des Batlapis, et d'autres idiomes encore.

Gagnons à présent la côte occidentale, la côte de l'Atlantique. Le domaine du système linguistique bantou est moins étendu ici que sur la côte de l'océan indien.

Au nord, il dépasse l'équateur de 4 ou 5 degrés, et confine aux langues des Nègres proprement dits.

La division septentrionale de ce groupe occidental comprend la langue de Fernando-Po, le *mpongoué*, le *dikélé*, l'*isoubou* et le *doualla*, le *congo*, qui de tous ces idiomes est le plus important, et quelques autres langues peu connues.

Plus au sud, entre autres idiomes, il faut distinguer le *bounda*, langue d'Angola, et le *héréro* parlé aux alentours du 19e degré de latitude méridionale. Ce dernier idiome confine, au sud, à un dialecte hottentot, le nama.

La classification de W Bleek est un peu différente. Il divise toutes ces langues en trois branches (1).

La première comprend le *cafir*, le *zoulou*, le *sétlapi*, le *sésouto*, le *tékéza*.

La seconde compte cinq subdivisions : tété, séna, makoua, kihiaou ; kikamba, kinika, kisouahili, kisambala ; bayéiyé (dans l'intérieur des terres) ; héréro, sindonga (langue des Ovambo), nano (dans le Bengouéla), angola ; congo, mpongoué.

La troisième comprend le dikélé, le benga (dans les

(1) Bleek. *A Comparative Grammar of South-Africa Languages*. Londres, 1869, p. 5.

îles de la baie de Corisco), le doualla, l'isoubou, la langue de Fernando-Po.

Il est assez difficile de se prononcer sur ce groupement. On ne connaît point toutes les langues du centre de l'Afrique méridionale ; de nouvelles découvertes, de nouvelles études aideront sans doute à classer d'une façon plus exacte les idiomes que l'on connaît déjà.

Il n'y a rien à dire de particulier des voyelles du groupe bantou, sinon qu'elles se prêtent volontiers à des contractions, à des suppressions euphoniques et à des variations assez nombreuses, mais toujours bien motivées. Les idiomes cafres sont plus raffinés en cela que beaucoup d'autres langues agglutinantes. On rencontre chez eux de véritables exemples d'harmonie vocalique, c'est-à-dire des exemples de la voyelle d'une syllabe s'assimilant à la voyelle d'une autre syllabe du même mot.

Le système des consonnes semble assez compliqué dans les différents idiomes du groupe bantou. Cela tient surtout à la grande quantité de consonnes doubles dont le premier élément est une nasale : *nt*, *nd*, *mp*, etc., etc.

D'autre part nous retrouvons ici une partie des « claquements », des consonnes « claquantes » dont nous avons parlé lorsqu'il s'est agi de la phonétique du hottentot. Les Cafres auraient emprunté aux Hottentots ces consonnes particulières ; en tout cas on ne les rencontre que dans les dialectes voisins du hottentot, par exemple dans les idiomes du rameau kafir-zoulou. Plus on s'éloigne de ce voisinage, moins ces consonnes deviennent fréquentes. Ainsi nous ne les trouvons pas en mpongoué. D'ailleurs, dans les idiomes cafres, ces claquements ne peuvent pas précéder d'autres consonnes (comme c'est le cas en hottentot) ; elles ne font que tenir la place d'autres consonnes. Des quatre claquements du hottentot il n'y en a que deux qui soient communément usités ici,

notamment le claquement dental. Des deux derniers l'un est fort rare, l'autre tout à fait inconnu.

Le nombre des autres consonnes est assez considérable. Elles sont soumises à des lois euphoniques, et les principes d'après lesquels elles correspondent les unes aux autres dans les différents idiomes sont des principes réguliers. Un grand nombre de ces concordances sont aujourd'hui connues et déterminées (1). Le cafir paraît plus avancé que ses congénères dans les voies de l'euphonie.

Les langues du système bantou ont ceci de particulier que le mot est formé en principe non point par des suffixes — c'est-à-dire par des éléments venant se placer après la racine, — mais bien par des préfixes, c'est-à-dire par des éléments placés en tête même de la racine. Si nous nous reportons à la théorie exposée ci-dessus, p. 58, nous voyons que la forme du mot cafir, tékéza, héréro, etc., est celle-ci : R R.

Parmi ces préfixes, les uns désignent le singulier, les autres le pluriel. En cafir, par exemple, les préfixes du singulier sont *ili*, *izi*, *u*, *ulu*, *um* ; ceux du pluriel sont *aba*, *ama*, *imi*, *izi*, *izim*, *izin*, *o*. Ainsi *umntu* veut dire « homme » et *abantu* « hommes » ; *udade* « sœur » et *odade* « sœurs ». Cela n'est qu'un exemple particulier, et les différents idiomes du groupe bantou n'ont pas tous aujourd'hui les mêmes préfixes formatifs ; mais ces préfixes d'apparence variée remontent cependant les uns et les autres à des formes communes plus anciennes. Il a existé, à une époque que nous ne pouvons déterminer, un idiome bantou commun ; cet idiome s'est divisé en diverses langues caractérisées les unes et les autres par des lois euphoniques particulières, et la forme des préfixes de cet ancien idiome s'est diversifiée naturellement dans les différentes langues auxquelles il donna naissance.

(1) Bleek. *Op. Cit.*, p. 81.

Nous venons de parler des préfixes *um*, *aba* et autres du cafir. La comparaison avec tous les autres idiomes du groupe bantou montre que la voyelle initiale de ces préfixes constitue réellement un autre préfixe. Les mots *umntu*, *abantu* se décomposeraient donc ainsi : *u-m-ntu*, *a-ba-ntu* et les éléments *m*, *ba* seraient (dans l'espèce présente) les vrais éléments dérivatifs du mot. Le sésouto (dialecte sétchouana) dit *motu* au singulier, *batu* au pluriel ; le séna *munnto* et *vanttu ;* le kihiaou (dialecte de Zanzibar) *mundu* et *vandu*. Mais en héréro nous retrouvons comme en cafir un autre élément préfixé : *omundu, ovandu ;* de même en congo *: omuntu, oantu.* Les auteurs qui se servent du mot « abantou » pour désigner l'ensemble de la famille, feraient donc mieux de s'en tenir simplement à celui de « bantou », qui est un dérivé de premier degré.

Voici d'ailleurs un tableau des formes de ce mot au singulier et au pluriel dans quelques-uns des idiomes qui nous occupent :

	Singulier.	Pluriel.
Kisouahili	*mtu*,	*watu*
Kinika.	*mutu*,	*atu.*
Kikamba.	*mundu*,	*andu.*
Kisambala	*muntu*,	*wantu*
Kihiaou	*mundu*,	*wandu.*
Séna	*munttu*,	*vanttu.*
Makoua.	*mûttu*,	*attu.*
Cafir	*umntu*,	*abantu.*
Zoulou	*umuntu*,	*abantu.*
Sétlapi	*mothu*,	*bathu.*
Sésouto	*motu*,	*batu.*
Tékéza.	*amuno*,	*vano.*
Héréro.	*omundu*,	*ovandu.*
Sindonga.	*umtu*,	*oantu.*
Nano.	*omuno*,	*omano.*
Angola.	*omutu*,	*oatu.*
Congo.	*omuntu*,	*oantu.*
Benga.	*moto*,	*bato.*
Doualla	*motu*,	*batu.*
Isoubon.	*motu*,	*batu.*

L'élément qui a pour mission d'indiquer la notion du cas se place également avant le nom. En héréro, par exemple, le signe de l'instrumental étant *na*, nous avons *nomundu* ou *namundu* « avec l'homme ». Il y a ici application d'une loi euphonique : la forme première était *naomundu* pour *na+omundu*. En cafir, où « homme » se dit *umntu* et « hommes » *abantu*, ainsi que nous l'avons vu, *ngomntu* veut dire « avec l'homme » et *ngabantu* « avec les hommes » : ici le signe de l'instrumental est *nga* (correspondant à *na* du héréro) ; nous voyons comment il se préfixe au mot formé par un premier élément dérivatif, soit singulier, soit pluriel.

Le nom adjectif se forme avec le même élément dérivatif que le nom substantif auquel il sert d'épithète ; s'il y a une différence, elle est au moins très petite. Le mot *kulu* signifiant « grand » en cafir, on dit *umntu omkulu* « homme grand », *abantu abakulu* « hommes grands ». Le mot *into* « chose » étant au pluriel *izinto*, on dit, dans cette même langue, *into enkulu* « chose grande », *izinto ezinkulu* « choses grandes ». En un mot, l'adjectif concorde forcément, quant à sa formation même, avec le mot substantif qu'il qualifie.

Dans une même phrase donc, le mot *kulu* « grand » pourra se voir juxtaposer quatre ou cinq préfixes différents, s'il est répété quatre ou cinq fois et sert d'épithète à autant de mots formés au moyen de tout autant de préfixes différents. Nous avons pris un exemple en cafir, nous eussions pu le prendre dans toute autre langue du groupe bantou. Le procédé est le même dans toutes ces langues ; de là les noms de langues allitérales, de langues concordantes, qu'on leur a donnés.

Le mécanisme de la façon d'exprimer les notions de temps et les notions de modalité, peut paraître assez compliqué, au premier abord, dans le système bantou. Au fond cependant, il n'en est rien. Ici, comme dans toutes

les langues agglutinantes, il n'y a qu'une simple agglomération de racines juxtaposées, une dérivation pure et simple.

La vraie caractéristique des langues appartenant à ce groupe, c'est la formation de ses mots au moyen de préfixes, d'éléments placés devant la racine ; c'est sur ce seul et unique point, qu'il était utile d'insister d'une façon particulière.

§ 5. Le poul.

Les Pouls ou Peuls, ou Foulas, occupent le centre de l'Afrique entre les 10e et 20e degrés de latitude ; à l'ouest, ils ne sont pas éloignés de la côte du Sénégal ; à l'est, ils s'étendent jusqu'au lac Tchad. C'est une région d'environ sept cent cinquante lieues de longueur, coupée à mi-chemin par le fleuve Niger. Sa largeur moyenne est d'environ cent vingt-cinq lieues, du 10e au 15e degré de latitude nord. Les dialectes principaux du poul sont le foutatoro, le foutadjallo, le boudou, le sokoto.

Le matériel phonétique du poul est peu compliqué ; on n'y rencontre point les gutturales de l'arabe en dehors des mots d'emprunt.

Le poul ne connaît pas la distinction du genre masculin et du genre féminin, mais il partage les êtres en deux catégories. Il distingue, d'une part, tout ce qui appartient à l'humanité : d'autre part, tout ce qui ne lui appartient pas : animaux et choses non animées. L. Faidherbe donne à ces deux genres les noms de *genre hominin* et *genre brute* (1).Cette distinction est capitale dans la grammaire poule. Les noms qui se rapportent à des

(1) Cf. Fr. Müller, *op. cit.*, t. III, p. 1.

êtres du genre hominin, substantifs, adjectifs ou participes, ont tous au singulier la terminaison *o*, qui n'est qu'une racine pronominale agglutinée : *gorko*, homme. La désinence du pluriel des noms du genre hominin est *bé*, qui n'est encore que le pronom *ils*, *elles*. S'agit-il du genre brute, la terminaison du singulier est une voyelle, ou bien *l*, ou bien *am* ; la désinence *o*, y est très rare ; le pluriel paraît des plus compliqués, et certaines lois euphoniques semblent jouer un très grand rôle dans l'agglutination des terminaisons au radical. Les consonnes initiales du mot au singulier peuvent se changer en d'autres consonnes quand le mot est au pluriel.

Le verbe est beaucoup plus simple. Les différents temps se forment, comme dans toutes les langues agglutinatives, par l'agglomération de divers éléments dont l'analyse demeure toujours assez claire.

La syntaxe du poul n'est pas compliquée. L'ordre même de la succession des idées détermine en principe l'ordre des mots dans la phrase. Ainsi, le nom du possesseur est précédé de celui de la chose possédée ; le régime direct ou indirect suit le verbe. Toute la difficulté du poul réside, en somme, dans la grande variété des principes euphoniques, mais c'est là une difficulté considérable.

En embrassant le mahométisme, les Pouls ont introduit dans leur langue un certain nombre de mots arabes, des termes religieux, des termes de droit, bien d'autres encore. Ce bagage mis en dehors de la question, il restait à savoir s'il y avait vraiment une parenté entre certaines langues du Sénégal, telles que le wolof et le sérère, et la langue poule, et, en cas d'affirmative, quelle pouvait bien être cette parenté. Que le wolof, le sérère, le poul, et d'autres langues encore, aient en commun un certain nombre de mots, c'est ce que l'on ne peut nier ; mais, dans l'état actuel des connaissances, il serait au moins téméraire d'établir sur une concordance lexique,

qui, en définitive, est faible, l'affirmation d'une prétendue parenté fort problématique. Faidherbe est très réservé sur cette prétendue alliance du poul, du wolof, du sérère; ce n'est pas sans de justes motifs. En théorie, elle n'est rien moins que prouvée. Nous savons que les Pouls n'ont atteint le Sénégal qu'après avoir traversé le centre de l'Afrique ; selon toute vraisemblance, c'est dans l'Afrique orientale qu'il faut chercher leurs anciens parents, c'est là que l'on peut trouver, s'il en existe encore, des idiomes alliés à leur propre langue.

§ 6. Les langues nubiennes.

Les ethnologistes rattachent les uns aux autres les Nubiens et les Pouls et en font une seule et même race, dont les premiers formeraient la division orientale, les seconds la division occidentale. En tout état de cause, les langues de ces deux populations paraissent différentes.

Le *nubien* proprement dit, l'idiome des Barabras, est parlé sur le cours du Nil, du 21e au 24e degré de latitude, par environ quarante mille individus.

Le *dongolavi*, parlé un peu plus au sud, en diffère assez peu.

Dans le sud du Kordofan (au nord du chilouk, langue d'un peuple nègre) est parlé le *toumalé*.

Le *koldadji* est un peu plus à l'ouest.

On rattache également au groupe nubien, mais cependant avec une certaine réserve, le *kondjara*, parlé dans une partie du Darfour et du Kordofan.

D'autres idiomes, enfin, pourraient être, eux aussi, apparentés à ce même groupe, mais les renseignements que l'on possède à leur sujet sont encore trop incomplets pour que l'on se prononce d'une façon définitive.

§ 7. Langues des Négritos.

On ne sait que bien peu de chose des idiomes parlés par les différents groupes de Négritos. Nous n'avons guère, pour l'instant, qu'à signaler l'existence de ces idiomes.

Le domaine de ces petits nègres, à tête plus ou moins arrondie, n'a aucune continuité. On en rencontre dans la presqu'île de Malacca (Sémangs), aux îles Andaman et Nicobar, aux Philippines (Hilloonas), dans les îles situées entre Mindanao et Luçon (Aétas). On en trouve encore en Papouasie, en Mélanésie (Hovelacque et G. Hervé, *Précis d'anthropologie*, p. 377). D'après quelques auteurs, ils auraient également formé le fond le plus ancien de la population de l'Inde et de celle de l'Indo-Chine. Quoi qu'il en soit, ce que l'on sait de leurs idiomes est fort peu de chose.

§ 8. Langues des Papous.

Le domaine ethnique des Papous est continu. Leur région principale est la Nouvelle-Guinée. C'est de là qu'ils se sont étendus sur Waïgiou, Misol, Céram, Bourou, une partie de Timor, Flores, Bathurst, etc. Vers l'est, ils occupent les îles Salomon, les Nouvelles-Hébrides, les Loyalty, la Nouvelle-Calédonie, les îles Viti. Mais dans toutes ces îles ne règne pas le même système de langues. Une partie des Papous, des Mélanésiens, parlent un idiome de la famille linguistique maléo-polynésienne. Gabelenz avait rattaché à cette famille le néo-calédonien, l'idiome des Loyalty ; on a reconnu qu'ils appartenaient

au groupe papou. (Voir au paragraphe des langues maléo-polynésiennes.)

On rencontre ici la dérivation par suffixe et celle par préfixe. Ainsi de *snûn* « l'homme », de *bien* « la femme », on forme *snûnsi* « les hommes », *biensi* « les femmes ». Les particules répondant à ce qu'on appelle les cas dans les langues à flexion, se placent ici devant le mot : *rosnûn* « de l'homme », *besnûn* « à l'homme », *rosnûnsi* « des hommes », *besnûnsi* « aux hommes ».

§ 9. Langues australiennes.

Les différentes langues de l'Australie (et elles sont nombreuses) paraissent toutes apparentées les unes aux autres, mais elles ne se rattachent à aucune autre famille linguistique.

Leur système phonétique est des plus simples ; il ne s'y rencontrerait ni sifflantes ni aspiration. Dans presque toutes ces langues, la notion même du nombre est peu développée ; quant à celle du genre, elle est totalement inconnue. Par contre, on trouve une certaine richesse dans le matériel des suffixes chargés de déterminer les relations du nom, ce qu'on appelle improprement les cas dans les langues agglutinantes.

On divise les langues australiennes en trois groupes. Celui de l'est, rapproché du grand Océan, est parlé dans une partie du Queensland et dans la Nouvelle-Galles du Sud. Il comprend le *kamilaroi*, ou kamilroi, près de la rivière Barwan ; le *koinberri ;* le *wiratouroi ;* le *wailwoun*, dans la région de la rivière Barwan, vers le fort Bourke ; le *kokai*, plus au nord, sur les rivières Maranoa et Kogoun ; le *wolaroi ;* le *pikoumboul ;* le *paiamba ;* le *kinki ;* le *dippil*, au nord de la baie de Moreton ; le

tourrouboul, près de la rivière de Brisbane. Le groupe central comprend les dialectes parlés au nord d'Adélaïde. Le groupe de l'ouest, enfin, comprend les dialectes parlés dans l'Australie occidentale du sud, à l'est et au sud de Perth.

Toutes ces idiomes appartiennent, comme on le voit, à l'Australie méridionale. On ne sait jusqu'ici que fort peu de chose (et parfois même l'on ne sait rien encore) des idiomes du centre et du nord de ce vaste territoire.

La phonétique des langues australiennes est fort peu compliquée. Elles ne connaissent qu'un petit nombre de voyelles et de consonnes ; il semble qu'elles ne possèdent point les explosives faibles (b, d, g). Elles ne forment les mots qu'au moyen de suffixes, jamais au moyen de préfixes, c'est-à-dire qu'elles placent l'élément dérivatif après le radical (comme le sanskrit, le latin, le grec) et non pas avant (comme les langues du système bantou). Les particules adjointes au nom pour rendre l'idée des différents cas sont fixées après lui : *tippin* « oiseau », *tippinko* « à l'oiseau » ; *punnul* « soleil » ; *punnulko* « au soleil ».

Le système de numération des Australiens est des plus restreints ; il compte jusqu'au nombre de quatre inclusivement, mais à partir de cinq il emploie une expression générale ayant le sens de grande quantité.

§ 10. Les langues maléo-polynésiennes.

On leur donne parfois le nom de langues océaniennes, bien qu'elles comprennent des langues parlées en Afrique (le malgache) et en Asie (la langue de l'île de Formose).

Disons d'abord que l'on divise en trois groupes principaux les langues maléo-polynésiennes : un groupe méla-

nésien (immédiatement à l'est des Papous et des Australiens) ; un groupe polynésien, à l'est du précédent ; un groupe malai, au sud-est de l'Asie.

Dans le GROUPE MÉLANÉSIEN (Codrington, *The Melanesian Languages*), on distingue les langues des îles Viti ; des Nouvelles-Hébrides (Annatom, Tana, Erromango, Mallikolo, etc) ; des îles Salomon ; les idiomes de la Micronésie : îles Carolines, Marshall, Gilbert. Le groupe linguistique, dit *mélanésien*, correspond donc, ethniquement, à une partie des Papous (Vitiens, etc.), qui ont abandonné leur ancien idiome, et à des proches parents des Polynésiens, à savoir les Micronésiens (Hovelacque et G. Hervé, *Précis d'anthropologie*, p. 781). Nous avons dit ci-dessus que le néo-calédonien et l'idiome des îles Loyalty appartenaient au domaine papou. En somme, il y a ici, comme souvent ailleurs, un départ important à faire entre la race et la langue.

Le GROUPE POLYNÉSIEN est bien plus étendu ; de la Nouvelle-Zélande (dans l'océan Austral) aux îles Sandwich (dans l'océan Boréal), il y a 60 degrés latitudinaux de différence et il y plus de distance encore entre la limite orientale et la limite occidentale du groupe qui nous occupe. Tout à l'est, nous trouvons la langue de l'île de Pâques ou Vaihou ; et en revenant vers l'ouest les langues des îles Gambier, de l'archipel Pomotou, des îles Marquises, des îles de la Société (Tahiti et autres), de l'archipel de Cook, des îles australes (Toubouaï et autres), des îles Tonga, des îles Samoa, des îles de l'Union, un peu plus au nord ; de Toucopia, limite occidentale du groupe polynésien ; puis, tout au nord, la langue des îles Sandwich (Havaï), et tout au sud, celle des Maoris, habitants de la Nouvelle-Zélande.

Le GROUPE MALAI comprend deux subdivisions : la branche du tagala et celle du maléo-javanais (Fr. Müller, *Allgemeine ethnographie*, p. 24). Dans le premier sous-

groupe, il faut compter : premièrement, les différents dialectes malais des îles Philippines, savoir : le *tagala*, à Luçon et dans les Philippines du nord ; le *bisaya*, dans les îles situées un peu plus au sud et dans le nord de Mindanao ; le *pampanga*, l'*iloko*, le *pangasinar*, tous parlés à Luçon (Cust, *A Sketch*, etc., p. 169) ; le *bicol*, la langue de l'île de Zébu ; — secondement, la langue des îles Mariannes ; — troisièmement, la langue de la partie orientale de l'île de Formose (au nord des Philippines et non loin de la Chine). La région occidentale appartient au chinois ; le domaine malai est le territoire montagneux de l'intérieur et de l'est. Ici, la population est peu et parfois point du tout cultivée ; on la rencontre çà et là à l'état sauvage. La preuve de l'antiquité de la migration malaie à Formose se trouve dans ce fait que le dialecte en question ne renferme point de mots hindous. Cette occupation a donc eu lieu avant les rapports des Hindous et des Malais; — quatrièmement, le *malgache*, à Madagascar, dépourvu également de tout élément emprunté aux Hindous, et dès lors importé depuis fort longtemps dans la région qu'il occupe.

Le second sous-groupe maléo-javanais est plus important. Tout d'abord, le *malai* proprement dit, parlé dans tout le centre de Sumatra, d'où il a rayonné, dans les îles voisines de Banca et de Billiton, sur toute la côte de Bornéo, dans la partie méridionale de la presqu'île de Malacca ; enfin, comme langue commerciale, dans presque toute la Malaisie et particulièrement aux îles Moluques. Le malai est, en effet, une sorte de langue internationale (les Malais sont essentiellement navigateurs et commerçants), qui est également comprise dans les relations commerciales sur les côtes de l'Inde, de Chine, de Siam et de Cochinchine. Les Hindous ont eu jadis une certaine influence sur la Malaisie ; ils ont donné au malai un certain nombre de mots tirés du vocabulaire aryen.

Plus tard, vers le douzième siècle, les Arabes, eux aussi, pénétrèrent en Malaisie ; ils y apportèrent l'islamisme, leur écriture, — si peu pratique, — mais on n'estime pas à plus de cent cinquante le nombre de mots arabes acceptés par le malai. Cet idiome est parlé par environ dix millions et demi d'individus ; sa littérature est considérable, mais n'a rien d'original. L'*atchinois*, parlé dans la partie septentrionale de Sumatra, par une population mahométane, est fort rapproché du malai ; quelques auteurs le regardent même comme un simple dialecte de ce dernier. Au sud de l'atchinois, toujours dans l'île de Sumatra et au nord du malai (qui, ainsi que nous l'avons dit, occupe la partie centrale de l'île), est parlé le *battak* par une population de soi-disant païens, cannibales, mais cependant assez civilisés et ayant leur écriture propre. (Au sud-ouest de l'île est parlé le « rédjang », au sud-est le « lampung », séparés tous deux du battak par le malai, qui détient le centre de Sumatra ; nous ne pouvons nous prononcer sur la nature de ces deux idiomes.) Le *javanais* est parlé par treize millions d'individus, non dans toute l'île de Java,mais seulement dans la partie centrale; idiome très important, dont la littérature a beaucoup emprunté à l'hindou et à l'arabe, et qui possède une écriture dérivée de l'indienne. Le brahmanisme et le bouddhisme ont été jadis introduits à Java ; au quinzième siècle, l'islamisme y pénétra, et il y domine aujourd'hui. Dans l'ouest de Java, c'est-à-dire à Batavia, est parlé le *sondéen*, qui ne diffère pas considérablement du dialecte précédent ; quatre millions d'individus, mahométans. Dans la partie orientale de l'île et dans l'île de Madura est parlé le *madurais*. Un peu plus à l'est le *balinais*, parlé par cinq cent mille individus, dont la religion est le brahmanisme ou le bouddisme. En remontant vers le nord, à Célèbes (à l'est immédiatement de Bornéo), on trouve au nord l'*alfourou*, qui se parle également dans

BIBLIOTHECA

certaines des îles Moluques ; au sud, le *boughi ;* à l'extrême sud-ouest, le *makassar.*

Nous avons dit que toute la côte de Bornéo était occupée par le malai. La population des Dayaks, qui habitent l'intérieur de l'île, est une population sauvage.

Quant aux langues de Sumbawa, de Sumba, de Floris, de Timor et des autres îles situées au sud-ouest du domaine malai, on n'a sur leur compte que des renseignements fort imparfaits, et il est encore impossible de les classer.

On trouvera dans le volume de R. N. Cust, *A Sketch of the modern Languages of East Indies*, des renseignements sur la distribution de ces différents dialectes.

Toutes les langues maléo-polynésiennes ont une origine commune ; elles sont indépendantes de toute autre famille linguistique. Bopp fit une tentative malheureuse pour les réunir aux langues indo-européennes ; d'autres auteurs voulurent les rattacher à une prétendue famille touranienne dont nous dirons quelques mots au paragraphe vingtième du présent chapitre. Ce fut peine perdue. Leur système phonétique est distinct et bien distinct de tous les autres, leurs racines sont parfaitement originales et ne se prêtent à aucun rapprochement avec les racines du système indo-européen ou de toute autre famille.

D'après Fréd. Müller, l'ancienne phonétique commune du maléo-polynésien se composait des trois explosives *k*, *t*, *p*, des trois nasales correspondantes, d'un *h*, d'un *r*, des sifflantes *s*, *f*, *v* et des voyelles *a*, *i*, *u* (prononcez « ou »), *e*, *o*. C'est plus tard seulement que parurent les autres sons que présentent les langues maléo-polynésiennes, par exemple *g*, *d*, *b*, *tch*, *dj*, *y*, *l*, etc. Les éléments qui s'adjoignent à la racine pour former les mots sont tantôt préfixés, tantôt suffixés ; parfois, dans certains

dialectes, on les intercale dans le corps même du mot : en d'autres termes, on les incorpore.

Des trois groupes maléo-polynésiens, le groupe malai présenterait les formes les plus pleines, les mieux développées ; la branche tagala se distinguerait particulièrement. Ensuite viendrait le groupe mélanésien, moins riche déjà ; enfin le groupe polynésien aurait considérablement à envier, sous ce rapport à la langue des Philippines, à celle de Formose, au malgache de Madagascar.

Serait-ce à dire que le groupe malai représenterait avec plus de fidélité les formes communes qui ont donné naissance au tahitien, au tagala, au javanais ? Cela, assurément, ne saurait être affirmé.

Ce qu'il convient de supposer, c'est que le polynésien a été détaché de l'ensemble de sa famille à une époque où la langue n'était pas encore fort développée, et que sa civilisation propre ne lui a pas permis de développer davantage. Ainsi que l'a très justement remarqué Fréd. Müller, les langues à flexion se divisèrent en familles distinctes à un moment où leur structure était déjà parfaite : à partir de l'époque où ces différentes familles se trouvèrent constituées, nous n'assistons plus au développement de leurs formes ; nous ne sommes plus témoins, au contraire, que de leur altération. Mais s'il s'agit d'autres langues que des langues à flexion, nous assistons à un spectacle différent. Ici, en effet, la séparation d'une seule et même famille en branches bien distinctes s'opère à un moment où la structure du système commun n'est pas encore parachevée ; chacun des différents idiomes, après s'être détaché de ses congénères, doit encore pourvoir par ses propres moyens à l'achèvement de sa structure propre et individuelle. Dans ces sortes de langues il est donc aisé de retrouver identité de racines et identité d'éléments servant à former les mots,

mais cela est tout et l'on ne peut espérer trouver identité de mots tout construits.

La grammaire maléo-polynésienne est celle de toutes les langues agglutinantes. Point de déclinaison véritable ; des particules font l'office de nos prépositions. Ainsi, le mot vitien *na tanoa* « le bassin » donne *ki na tanoa* « au bassin », *î na tanoa* « dans le bassin », *kei na tanoa* « avec le bassin » ; le polynésien de la Nouvelle-Zélande dit *te tanata* « l'homme », *a te tanata* « de l'homme », *ki te tanata* « à l'homme » ; même procédé en malai : *ânak* « l'enfant », *âkan ânak* « à l'enfant », *deri ânak* « [venant] de l'enfant ».

Aucun élément particulier ne vient s'agglutiner au nom pour former le pluriel. Dans la langue du Viti, par exemple, *a tamata*, signifie aussi bien « les hommes » que « l'homme » ; dans celle d'Erromango, *niteni* signifie aussi bien « les fils » que « le fils ». C'est à l'aide de procédés en quelque sorte artificiels que l'on arrive à indiquer la pluralité. Tantôt c'est en opposant à un article indiquant le singulier, un article différent indiquant le pluriel : maori, *te tanata* « l'homme », *na tanata* « les hommes », *ki te tanata* « à l'homme », *ki na tanata* « aux hommes » ; tantôt c'est en employant un mot signifiant accumulation, foule : tahitien, *te taata* « l'homme », *te mau taata* « les hommes ».

Le procédé est analogue, tout aussi artificiel, tout aussi primitif, dans les différents idiomes du groupe malai : ou bien l'on adjoint au mot dont il s'agit un autre mot exprimant le grand nombre, la collectivité, ou bien l'on redouble le mot. Ce redoublement a ses règles particulières. C'est ainsi que dans la langue de Formose, on redouble la première syllabe : *sjien* « la dent », *sisjien* « les dents » ; en javanais, tout le mot peut être redoublé : *ratu* « le prince », *raturatu* « les princes ».

Le genre n'est point désigné davantage par la juxta-

position au mot d'un nouvel élément. On se sert d'un mot accessoire. Dans la langue de Viti, par exemple, c'est *tagane* « masculin » et *aleva* « féminin » ; *a gone tagane* « garçon », *a gone aleva* « fille ». En tahitien, *metua* signifie « parent », et se rapporte aussi bien au père qu'à la mère ; mais s'il s'agit de désigner tout spécialement le père ou la mère, on ajoute à ce mot ou bien celui de *tane*, où bien celui de *vahine*. Pour les animaux, on se sert de deux autres expressions : *oni* et *ufa ;* l'on dit, par exemple, *moa oni* « coq » *moa ufa* « poule » (1).

D'autre part, point de conjugaison véritable : c'est par des particules, par des affixes, par l'emploi de mots auxquels on ne donne plus qu'une signification toute subordonnée, que l'on arrive à donner au mot principal l'idée secondaire de temps ou de mode. En général, le mot principal, le verbe de cette soi-disant conjugaison, est placé à la fin ; ainsi, dans la langue mélanésienne d'Annatom, l'on dit : *ek asaig* « je dis », *ek mun asaig* « j'ai dit », *ekis asaig* « je disais », *ekis mun asaig* « j'avais dit », *ekpu asaig* « je dirai », *eku vit asaig* « si je dis », et ainsi de suite. Cela toutefois est loin d'être une règle absolue.

Le fonds de la numération est décimal.

Nous avons dit tout à l'heure que dans les langues maléo-polynésiennes les éléments qui venaient se juxtaposer au mot pour le dériver, pour lui donner un sens secondaire, pouvaient être placés en tête de ce mot (comme c'est le cas régulier dans les langues du système bantou), ou à la fin du mot (comme c'est le cas dans les langues indo-européennes et les langues australiennes), ou, enfin, être incorporés dans le mot.

Lorsque les divers idiomes polynésiens forment leurs causatifs en dérivant le mot principal par *faka, haa,*

(1) Gaussin, *Du dialecte de Tahiti, de celui des îles Marquises et, en général, de la langue polynésienne.* Paris, 1853.

fa'a, *vaka*, etc. (maori : *kite* « voir », *vakakite* « fait voir » ; tahitien : *mate* « mourir », *haamate* « tuer », il y a préfixation. Il en est de même encore dans les mots malais *berpâkei* « vêtu, muni d'un vêtement », *berbîni* « marié, pourvu d'une femme », tirés de *pâkei* « vêtement », *bîni* « femme ».

Nous trouvons, au contraire, que la dérivation est opérée au moyen d'un suffixe, c'est-à-dire d'un élément placé après le radical, dans le tagala *putian* « blancheur », provenant de *puti* « blanc », *bigayan* « don », provenant de *bigay* « donner ». Il en est de même dans le maori *korerotia* « dit », dérivé de *korero* « dire », dans *kaina* « nourriture »,dérivé de *kai* « manger ». C'est ainsi que le polynésien forme son passif.

Enfin, dans le groupe malai, nous trouvons parfois l'élément dérivatif incorporé, c'est-à-dire placé à l'intérieur même du radical. Ainsi, en tagala, de *tapay* « pétrir » est dérivé de *t-in-apay* « pain », de *sipit* « saisir » est dérivé de *s-in-ipit* « ancre ». Plusieurs éléments dérivatifs peuvent se trouver dans un seul et même mot.

Les langues maléo-polynésiennes ont presque toutes une littérature plus ou moins développée. Chez les Polynésiens, on trouve un grand nombre de contes, de récits et de chants traditionnels. La littérature du malai est même assez riche. A la vérité, elle a beaucoup emprunté; ses écrits philosophiques sont inspirés par ceux des Hindous et des Musulmans ; mais ses romans et ses contes lui appartiennent souvent en propre, et ses poésies sont presque fugitives, des dialogues, des dictons, des fables, mais encore de vrais poëmes épiques et dramatiques. Le javanais possède une littérature qui doit beaucoup au sanskrit, non-seulement dans ses allures et son esprit général, mais encore dans son vocabulaire. Il a toutefois, lui aussi, ses poëmes et ses chants originaux, ses fables et ses légendes.

Le malai s'écrit avec les caractères arabes, que l'islamisme lui a fait connaître ; l'on ignore quel était son ancien alphabet. Les autres idiomes du groupe malai (tagala, javanais, makassar, etc.) ont emprunté leurs différents systèmes à un vieil alphabet hindou.

§ 11. Le japonais.

L'on a souvent cherché à rapprocher le japonais des langues ouralo-altaïques : du mongol, du turc, du magyar, du suomi et de tous les autres idiomes de cette famille.

(Voir Pott, *Ueber die Verschiedenheit des meuschlichen Sprachbaues* [Humbolt], t. Ier, p. CCCLXXIII ; Boller, *Acad. des sciences de Vienne*, t. XXIII.)

C'est du continent asiatique sans doute que sont venus les Japonais dans les îles qu'ils occupent aujourd'hui ; mais s'ensuit-il que leur langue ait une origine commune avec les langues du continent, même les plus voisines ? En aucune façon : il ne suffit point, pour établir ce prétendu fait, d'une simple et gratuite assertion.

Jusqu'à présent, on n'a donné aucune preuve sérieuse de cette prétendue parenté. On a bien dressé des listes de centaines de mots qui semblent offrir entre eux plus ou moins d'analogie, mais il ne s'agit pas ici de faire des étymologies. Les cinq cents homophones mongolo-japonais, que l'on s'est plu à découvrir, ne font pas avancer la question d'un seul pas. Autant comparer entre eux l'article portugais *o*, *a*, l'article magyar *a*, et l'article basque *a*. Cela n'est pas sérieux. Veut-on arguer du grand nombre des soi-disant concordances de mots japonais et de mots mongols ou magyars, l'on ne fait

qu'aggraver un cas déjà détestable : plus on entasse de semblables fantaisies, moins l'on devient excusable.

C'est en vain également que l'on invoque telles ou telles analogies dans la syntaxe : le bulgare qui place après le substantif l'article qu'il s'est fabriqué, serait-il, pour cela, allié au roumain, au basque, chez lesquels l'article est suffixé au nom ? C'est faire preuve, derechef, d'une profonde ignorance de la méthode linguistique que de demander à la syntaxe, dont les lois sont toutes secondaires, la raison du plus ou moins d'affinité des langues. En dehors d'une communauté de racines, il n'y a rien à espérer d'où puisse venir quelque preuve sérieuse relativement à cette question de l'origine unique ou multiple de deux ou plusieurs idiomes. Les prétendues similitudes de syntaxe n'ont pas plus de valeur que les amas de comparaisons de mots tout faits ; encore un coup, plus on se plaît à en entasser, moins on fait preuve d'esprit critique.

Il se peut que la langue japonaise soit rattachée plus tard à quelqu'une des familles linguistiques actuellement reconnues ; mais, dans l'état actuel des connaissances, il est impossible de rien avancer de certain ; il se peut que les recherches aboutissent dans ce sens, mais elles n'ont pas encore abouti.

Le japonais occupe la partie sud et centrale de l'archipel situé entre la mer du Japon, ainsi que la mer Bleue, et le Pacifique ; il comprend un certain nombre de dialec es qui ne paraissent pas éloignés les uns des autres.

L'écriture japonaise actuelle est assez difficile. Elle dérive des caractères chinois et remonte aux premiers siècles de notre ère, vraisemblablement au troisième. Chose assez singulière, cette écriture idéographique aurait été substituée à une écriture alphabétique plus ancienne, empruntée aux Coréens. De même que dans le système chinois, les signes, en japonais, s'écrivent de

haut en bas en colonnes parallèles, dont la première est celle de droite. Outre l'écriture cursive appelée *hirakana*, qui est universellement répandue dans le pays, il y a un système particulier, nommé *katakana*, dont les signes sont incontestablement plus simples, mais qui n'est guère employé que par les étrangers peu familiarisés avec l'autre système.

Notons d'ailleurs que l'écriture japonaise est une écriture syllabique. Abstraction faite des voyelles (*a*, *i*, *u*, *e*, *o*), elle procède par syllabes, par groupes d'une consonne accompagnée d'une voyelle : *ka*, *ki*, *ko*, etc. Le syllabaire japonais actuel comprend soixante-douze signes ; il en a compté jadis un tiers de moins environ.

Une nouvelle évolution semble devoir se produire un jour ou l'autre dans la transcription de la langue japonaise ; cette réforme considérable et très justifiée serait l'adoption de l'alphabet latin.

La première assemblée du Congrès des orientalistes a mis cette question à l'étude, et nous pensons qu'il y a quelque chance de succès (1).

On a pu constater une fois de plus, lors de cette tentative, à quel point il serait avantageux d'introduire dans nos établissements typographiques quelques caractères nouveaux et peu compliqués destinés à préserver des plus dangereuses confusions dans la transcription des langues qui ne se servent pas des caractères latins. Notre *ch*, par exemple, qui est le *sh* anglais, le *sch* allemand, le *sz* polonais, le *s* magyar, demande à être rendu par un signe unique lorsqu'il s'agit de la transcription d'un texte écrit en caractères particuliers : ce sera évidemment le signe *s* qu'emploient les Croates et les Tchèques. L'on rendra par le signe croate *ž* notre *j* français, qui, en allemand et dans les langues slaves pour les-

(1) *Congrès des orientalistes*. Paris, 1873.

quelles on emploie l'écriture latine, a le son de notre « y » (1). Sans prétendre atteindre à une simplification parfaite, l'on peut au moins s'arrêter à un système uniforme. La langue japonaise s'y prêterait sans difficulté.

La phonétique du japonais est assez simple. La formation des mots se prête à démontrer clairement ce que c'est qu'une langue agglutinative. La notion des cas est rendue très distinctement par le fait que des racines secondaires qui ont perdu leur indépendance et n'indiquent plus qu'une idée de relation, viennent se suffixer, s'annexer à la racine principale.

Quelques auteurs voudraient, dans la transcription latine des textes japonais, séparer par un trait d'union, par un tiret, le thème du mot d'avec ces éléments de rapport juxtaposés : *hito-no* « de l'homme », *hito-de-* « avec l'homme ». Mais ce système ne saurait être défendu par aucune bonne raison. Autant voudrait séparer en français le signe du pluriel *s* d'avec le reste du mot et écrire, par exemple, « père-s, dame-s, bon-s », ce que l'on justifierait par des motifs tout aussi valables. La juxtaposition la plus intime est le propre même des langues agglutinatives, et l'on ne pourrait, sans négliger le mode très caractéristique de la formation du mot dans ces sortes de langues, les représenter dans l'écriture tout autrement qu'elles n'existent en réalité dans la parole. Tout au plus pourrions-nous admettre que l'on réunît par un tiret au mot principal les préfixes *o*, *me*, qui déterminent le genre : *neko* « chat » : *o-neko* « matou », *me-neko* « chatte ». Quant aux signes du pluriel (tels que *tatsi*) ils doivent suivre, tout comme les éléments indiquant les cas, le principe de juxtaposition pure et simple : *hitotatsino* « des hommes », *hitotatside* « avec les hommes », au singulier *hitono*, *hitode*.

(1) E. Picot. *Tableau phonétique des principales langues usuelles*. Revue de linguistique, t. VI, p. 362.

S'agit-il de rendre les notions de temps et de mode, le japonais admet, comme toutes les autres langues agglutinantes, ces séries d'éléments juxtaposés les uns aux autres, dont nous avons déjà parlé et qui déterminent d'une façon de plus en plus précise le sens de la racine principale : élément négatif, élément causatif, élément optatif, et ainsi de suite. Il nous semble peu utile de dresser une liste d'exemples faciles à trouver dans les grammaires japonaises. Ces exemples seraient absolument analogues à ceux que nous avons déjà cités et que nous aurons à reproduire en parlant avec un peu plus de détails d'autres langues agglutinantes.

Sauf quelques publications éparses, la littérature japonaise n'a pas encore trouvé d'historien ; son intérêt pourtant est manifeste. L'histoire, le roman historique, les contes et les romans y tiennent une place considérable. Les ouvrages japonais de philosophie religieuse et de poésie sont également en grand nombre, et dans l'ordre des sciences, ceux de philologie et de botanique ont aussi leur importance. Ce ne sera pas sans doute une tâche très aisée que de discerner dans toutes ces compositions la part réellement japonaise de la part due à l'influence chinoise qui s'est fait sentir, notamment vers le troisième siècle de notre ère ; mais on peut prévoir que ce travail plein d'intérêt ne demeurera pas longtemps à l'état de desideratum.

Les mots chinois qui se sont introduits dans la langue japonaise avec cette influence littéraire sont soumis, tout comme les autres, au principe de juxtaposition ; c'est ainsi qu'en français nous mettons au pluriel en *s* des mots empruntés au langues germaniques : « meurtres, heaumes », ou aux langues slaves : « cravates, verstes ».

§ 12. Le coréen.

Cette langue a été rattachée à différents idiomes agglutinants, notamment au japonais. Sans nier absolument la possibilité du fait, nous attendons, avant de l'accepter, qu'on veuille bien l'appuyer de quelques arguments sérieux. Jusqu'à ce jour on ne s'est guère contenté que d'assertions à peu près gratuites.

De tous les idiomes de l'extrême Orient, le coréen d'ailleurs est le moins connu et le moins étudié. Il possède un véritable alphabet composé de voyelles et de consonnes figurées individuellement ; son écriture, en d'autres termes, est alphabétique. Cet alphabet assez simple daterait du quatrième siècle de notre ère ; mais son origine, en dépit de toutes les suppositions faites à ce sujet, est encore inconnue.

En coréen, de même que dans les autres langues agglutinantes, des postpositions viennent se joindre intimement au mot pour rendre les différentes idées de rapport, de relation, que les langues à flexion expriment par leurs cas. Le pluriel s'indique par la répétition du mot ou l'adjonction d'un autre mot dont le sens est celui de « tous » ou « beaucoup ».

Dans le lexique coréen, il s'est introduit un très grand nombre de mots chinois que l'on peut reconnaître sans trop de difficulté, bien que leur mode de prononciation soit assez varié.

§ 13. Les langues du centre et du sud de l'Inde.

L'Inde centrale et méridionale est peuplée par des Kols et des Dravidiens, peuples de peau foncée. Les premiers sont disséminés, sous différents noms, et par groupes plus ou moins importants, depuis le territoire de Nagpour, à l'ouest, jusqu'au Gange inférieur, à l'est. L'on ignore quelle peut bien être leur origine. Tandis que les Kols de l'ouest ont reçu et parlent un idiome aryen (soit l'hindi, soit le marathi), ceux de l'est, les Kolarians, ont conservé leur propre et vieil idiome, le kourkou, le korwa, le moundari, le santhal, le bhoumidj, le ho, le djouang.

La phonétique de ces langues est riche. Elles dérivent les mots au moyens de suffixes, parfois au moyen d'infixes : en santhal, *dapal* « couvrir », *danapal* « couverture » ; en moundari, *dub* « s'asseoir », *dunub* « siège ». Le système numérique est vigésimal.

D'après Cust, ces différents idiomes seraient parlés par environ deux millions d'individus.

Les langues dravidiennes, que l'on a appelées également *langues tamouliques*, *langues tamiliennes*, *langues malabarres*, tirent leur nom d'un mot hindou. Ce mot servait primitivement aux brahmanes à désigner ceux d'entre eux qui s'étaient établis dans cette partie de l'Inde qu'on appela plus tard le Décan ; il ne tarda pas à devenir le nom même de cette contrée et s'appliqua plus spécialement à la région où se parlait le tamoul, la plus importante des langues dravidiennes.

Ces langues occupent toute la partie méridionale de la péninsule cisgangétique, depuis les monts Vindhya et

la rivière Narmadâ (les Anglais écrivent Nerbuddha) jusqu'au cap Comorin. Dans cette vaste région, peuplée d'environ cinquante millions d'habitants, on trouve quelques colonies européennes ou musulmanes, mais le nombre des indigènes qui se servent exclusivement des idiomes dravidiens peut être évalué à quarante-cinq millions environ.

R. Caldwell, dans son important ouvrage sur les langues dravidiennes, les divise en deux groupes — nous ne disons pas en deux familles — selon qu'elles sont ou ne sont point cultivées. Le premier groupe comprend six langues : le *tamoul*, le *malayâla*, le *télinga*, le *kanara*, le *toulou*, le *koudagou ;* le second en comprend six ou huit dont nous donnerons tout à l'heure l'énumération.

Par la richesse de son vocabulaire, aussi bien que par la pureté et l'ancienneté de ses formes, le *tamoul* ou tamil joue dans la famille dravidienne le rôle que joue le sanskrit dans l'ensemble des langues qui lui sont apparentées. Le tamoul est la langue usuelle des quinze millions d'individus qui habitent toute la plaine à l'est des monts Ghattes, depuis Paliacate (un peu au nord de Madras) jusqu'au cap Comorin, et le sud de la côte occidentale jusqu'à Trivandrum. Il s'est étendu également sur le nord de l'île de Ceylan. On parle tamoul à Madras, Pondichéry, Karikal. Le *malayâla* est parlé par plus de trois millions et demi d'individus le long de la côte malabare, dans la longue bande de terre qui s'étend entre les Ghattes, à l'est, et le golfe Persique, à l'ouest, de Trivandrum à Mangalore. On regarde le malayâla comme un très ancien dialecte du tamoul, où les mots d'origine hindoue se sont introduits en assez grand nombre ; en fait, ces deux dialectes sont aujourd'hui parfaitement distincts. Le *toulou* ou toulouva, répandu jadis au nord du malayâla, est confiné actuellement aux environs de Mangalore, à l'est des Ghattes,

et le nombre de ceux qui le parlent ne dépasse pas de beaucoup celui de trois cent mille. Evidemment ce dialecte est destiné à périr dans un avenir peu éloigné, et son territoire est fortement pénétré par les idiomes qui l'avoisinent. On l'a pris parfois pour un dialecte du malayâla ; il en diffère pourtant d'une façon assez tranchée et constitue en réalité une véritable branche de la famille dravidienne. Le *kanara* ou kannada, pour parler plus exactement, occupe le nord du pays dravidien ; il s'étend sur le plateau de Mysore, sur la partie occidentale du territoire de Nizam, et est parlé par plus de neuf millions d'individus. Cette langue est d'un haut intérêt ; elle a conservé en maintes circonstances des formes très anciennes et très pures, plus anciennes parfois que les formes mêmes du tamoul. Le *télougou*, ou télinga, ou ténougou, est l' « ândhra » des écrivains hindous. Au nord du télinga, sont parlés le khond et l'oraon, dont nous dirons quelques mots tout à l'heure ; au sud, le tamoul. On parle télinga à Tchicacole, Yanaon, Haïderabad ; le territoire télinga comprend environ seize millions et demi d'individus et surpasse sous ce rapport tous ses congénères, même le tamoul ; mais il leur cède d'une façon très manifeste s'il est question de la bonne conservation des formes grammaticales. Sa phonétique aussi a beaucoup varié ; elle est d'ailleurs très harmonieuse et le télougou a reçu à juste titre le nom d'italien du Décan. De tous les idiomes dravidiens que l'on peut dire idiomes cultivés, le *koudagou* est le moins important ; il est parlé par cent cinquante mille individus environ, à l'ouest de Mysore. Après l'avoir regardé précédemment comme un dialecte du kanara, R. Caldwell s'est décidé, dans la seconde édition de son livre, à lui donner une place indépendante.

Les idiomes secondaires, ceux qui n'ont jamais été écrits, sont comme nous l'avons dit plus haut, au nombre

de six ou huit. Le *kôta* est parlé par onze cents Indiens presque sauvages, qui habitent une des gorges des Nilgherries ; il se rapproche beaucoup du kanara. Le *touda*, ou toda, est également l'idiome d'une tribu des Nilgherries ; d'après les derniers renseignements on ne compterait pas plus de sept cent cinquante Toudas. Le *gônd*, au contraire, est parlé par plus d'un million six cent mille individus ; c'est la langue de la partie montagneuse des territoires de Gôndvâna, de Nagpour, de Saugor et de la Nerbudda. Le *khond*, ou kou, est usité à Goûmsour, sur les frontières d'Orissa et dans la partie orientale du Gôndvâna ; on évalue à deux cent soixante-dix mille le nombre des individus chez lesquels il est usité. Le *râdjmahâl*, ou mâler, et l'*orâon* sont parlés dans l'Inde centrale, le premier par quarante mille, le second par plus de deux cent soixante mille individus ; ces deux derniers dialectes sont assez rapprochés l'un de l'autre.

Quelques auteurs ajoutent à cette dernière liste l'idiome *badaga*, usité dans une partie des Nilgherries, mais R. Caldwell le regarde comme un vieux dialecte du kanara qui ne possède aucun titre à être classé à part.

Quatre des établissements français de l'Inde sont situés en pays dravidien. Les deux plus importants, Pondichéry et Karikal, sont en pays tamoul. Mahé est sur la côte où l'on parle malayâla ; Yanaon enfin est dans la région du télinga.

Nous avons tracé une esquisse rapide des limites dans lesquelles les langues dravidiennes se sont maintenues depuis les temps historiques, subissant l'influence (pourtant assez lente) des idiomes hindous, mais résistant avec énergie à l'invasion musulmane et à la civilisation anglaise. Ces langues, si vivaces aujourd'hui encore, occupaient-elles jadis une région plus étendue ? Faut-il penser notamment qu'elles ont été refoulées dans leurs

limites actuelles par les premières immigrations aryennes ? Le fait est vraisemblable, probable même ; mais l'on n'en a donné jusqu'à ce jour aucune preuve démonstrative. On a supposé seulement que les éléments étrangers des dialectes de l'Inde septentrionale pouvaient avoir une origine dravidienne ; mais, outre qu'ils sont peu nombreux et de peu d'importance, il est fort difficile, non seulement de les analyser, mais encore de les déterminer. Dans la famille dravidienne elle-même, une grande partie du vocabulaire de certains idiomes incultes est de provenance inconnue. Il faut donc, pensons-nous, n'accepter qu'avec une grande réserve tout ce que l'on peut dire de l'ancienne extension des langues dravidiennes.

On peut avancer au moins, en toute sûreté, qu'elles ne se rattachent à aucune autre famille linguistique et qu'elles forment un groupe tout à fait indépendant. Tour à tour on en a fait des langues scythiques — ce qui, soit dit en passant, ne signifie absolument rien, ainsi que nous le verrons en parlant plus loin de la prétendue langue scythique ; — des langues affiliées au groupe ouralo-altaïque, au groupe indo-européen, au groupe sémitique, à bien d'autres groupes encore. Tous les rapprochements établis à ce sujet ont péché par un manque absolu de méthode. On a comparé des mots tamouls, des mots télougous, à des mots sanskrits, à des mots hébreux, à des mots choisis et pris dans toutes sortes de langues ; c'est le procédé habituel des personnes qui prétendent apparenter les langues au moyen de l'étymologie. Ce n'est point le tamoul ou le télougou qu'il faut comparer au sanskrit ou à l'hébreu : il s'agit de restituer avant tout la langue commune dravidienne, et c'est de la comparaison de ce type général avec le type des autres familles linguistiques que pourrait se dégager la réponse favorable à une prétendue communauté d'origine. Répétons cependant que ce qui est acquis d'ores et déjà paraît plus que

suffisant pour appuyer la thèse d'une complète indépendance des langues dravidiennes.

Bien avant que l'on ne connût le sanskrit, on s'occupait des langues dravidiennes. Elles avaient été découvertes de bonne heure par les navigateurs hollandais, danois, français et anglais.

Ces derniers s'empressèrent de les apprendre, dans l'intérêt de leur commerce d'abord, puis dans un but de propagande religieuse.

Les missionnaires composèrent les premiers nombres de grammaires et de vocabulaires dont la plupart n'ont jamais vu le jour.

J. Vinson a publié dans la *Revue de linguistique* nombre d'articles sur les langues dravidiennes. L'important ouvrage de Caldwell (*A comparative Grammar of the Dravidian or South Indian Family of Languages*) est devenu à juste titre un livre en quelque sorte classique, bien qu'il sacrifie à la fallacieuse théorie du touranisme et à la théorie non moins fallacieuse d'une communauté originelle des langues.

On trouve une carte des langues de l'Inde dans J. Beames, *Outlines of Indian Philology*, et dans R. N. Cust, *A Sketch of the modern Languages of the East Indies*.

La grammaire dravidienne est d'une grande simplicité. Nous allons essayer d'exposer en quelques pages l'ensemble de ses éléments et de ses procédés, sans entrer dans des détails trop particuliers.

La phonétique ne présente point de difficultés sérieuses ; son matériel est assez restreint. On ne compte dans les langues dravidiennes littéraires que les voyelles *a*, *e*, *i*, *o*, *u* (« ou » français), tantôt brèves, tantôt longues, et les deux diphthongues *ai* (que dans certains cas l'on prononce *ei*) et *au*. Cette dernière, au moins, n'appartenait pas au type dravidien commun. Par la suite des temps ces voyelles se sont affaiblies, leur prononciation

s'est atténuée, et il est résulté de ce fait un certain nombre de sons nouveaux, intermédiaires entre les différentes voyelles fondamentales ; ces différents sons se laissent parfaitement distinguer, mais dans l'écriture on ne les figure point. Sous ce rapport, le tamoul vulgaire diffère très sensiblement du tamoul littéraire, du tamoul écrit.

Les consonnes, elles aussi, sont en petit nombre dans les idiomes dravidiens. On compte cinq groupes d'explosives fortes et faibles (gutturales, palatales, linguales, dentales, labiales), et chacun de ces groupes possède une nasale de son ordre, soit quatre sortes de *n* et une *m*. Ajoutons *y*, *r*, *l*, *v*, un *r* fort ou double, deux continues linguales et une seule sifflante, *ç*.

Il existe également une nouvelle classe d'explosives particulières au tamoul et au télinga ; d'après R. Caldwell, leur prononciation serait celle de « tr, dr » ; pour J. Vinson, il ne s'agirait ici que « t » et de « d » mouillés au commencement même de leur émission. Ces deux auteurs entendent d'une façon différente un seul et même son ; cela prouve au moins qu'il est bien particulier. J. Vinson rend ces consonnes par *t'*, *d'*.

L'aspiration est inconnue aux langues dravidiennes. Leur système de consonnes devait être autrefois bien plus simple encore qu'il ne l'est aujourd'hui. J. Vinson pense, par exemple, que les « tch » et « dj » que l'on y rencontre sont d'origine relativement récente. Ces consonnes d'ailleurs, tout comme les voyelles, ont été altérées dans la prononciation populaire. En tamoul et en malayâla les dentales ont aujourd'hui une tendance marquée vers le *th* anglais doux ; en télinga le « tch » et le « dj » deviennent parfois « ts » et dz »..

La prononciation de ces différents sons est d'ailleurs assez facile. Les seules consonnes qui puissent nous sembler un peu étranges sont les linguales, que l'on

appelle habituellement, mais à tort, consonnes cérébrales. L' *l* des finales anglaises en *ble* donne une idée approximative de ce que sont ces consonnes linguales. Les langues dravidiennes en connaissent cinq : un *t*, un *d*, un *n*, un *j* ou *r*, un *l*. On les transcrit dans l'alphabet latin par les lettres *t*, *d*, etc., munies d'un point en dessous. Le sanskrit possède aussi des consonnes linguales, mais chez lui elles ne sont point organiques , elles paraissent constituer au contraire un caractère particulièrement distinctif des idiomes dravidiens.

Quant aux lois phonétiques qui peuvent se dégager de la comparaison de ces différents idiomes et de leurs variétés dialectales, nous n'en signalerons qu'une seule, qui est également familière aux langues indo-européennes : le *k* du kanara correspond souvent à un *tch* en télinga, à un *ç* en tamoul ; ainsi le mot « oreille », qui est *çévi* dans cette dernière langue et *tchévi* en télinga, se prononce *kévi* en kanara, et telle devait être la forme primitive.

Deux autres faits intéressants sont propres aux langues dravidiennes. Au commencement des mots la consonne *r* est proscrite ; s'agit-il d'un mot emprunté à une langue étrangère et commençant par *r*, on fait précéder cette consonne d'une voyelle : ainsi le mot sanskrit *râjâ* est représenté en tamoul par *irâyan* ou *irâçan*. Le second fait est plus curieux. Aucun mot ne peut commencer par une explosive douce (b, d, etc.), et aucune explosive dure (p, t, etc.) ne peut se trouver seule, isolée, dans le corps même d'un mot ; le tamoul empruntant au sanskrit le mot *gati*, le rend donc par *kadi* en se conformant à cette double règle. Les lois phonétiques des idiomes dravidiens n'ont pas encore été suffisamment étudiées pour que l'on puisse établir d'une façon définitive les principes qui président à la formation des mots. On en sait cependant assez pour classer

entre eux le tamoul, le télinga et leurs congénères et pour être fixé sur leur âge relatif. Les mots dravidiens paraissent se ramener à des racines dissyllabiques, et en comparant entre elles ces diverses racines, ces différents radicaux, on les ramène à leur tour à des éléments plus anciens. A vrai dire, cette étude n'est qu'ébauchée, elle est à peine commencée, mais on peut prévoir déjà que le monosyllabisme primitif des racines dravidiennes sera dûment établi par les travaux ultérieurs.

La dérivation dans les langues dravidiennes est nettement agglutinante. Elle s'opère toujours par la suffixation d'éléments nouveaux. Ainsi, à une racine comportant l'idée générale du mot, on ajoutera un élément chargé d'indiquer que l'action a lieu présentement, puis un élément comportant l'idée de négation, puis enfin un autre élément désignant la personne, et le résultat de cette agrégation, de cette agglomération sera un mot signifiant par exemple, « tu ne vois pas » et qui se décompose ainsi : « voir + présentement + non + tu ». Notons-le bien, le sens de chacun de ces éléments est toujours présent à l'esprit des Dravidiens : ils les traitent de la même façon que nous traitons, nous, nos pronoms, nos articles, nos prépositions.

A la vérité, un grand nombre de ces mots formatifs ont été tellement altérés par la suite des âges, que leur figure primitive est devenue méconnaissable ; mais beaucoup d'entre eux sont encore en usage dans le langage courant, avec leur sens naturel de demeure, contact, voisinage, conséquence, etc., etc. Ajoutons que plusieurs de ces éléments dérivatifs changent de l'une des langues congénères à l'autre, ce qui prouve bien l'indépendance originelle de ces suffixes.

Si l'on reconnaît sans peine combien les langues de cette espèce ont d'avantage sur les idiomes purement monosyllabiques, chez lesquels les racines ne se subor-

donnent que très-imparfaitement les unes aux autres (ainsi que nous l'avons vu dans notre troisième chapitre), d'autre part il est aisé de comprendre à quel point les idiomes à flexion leur sont supérieurs lorsqu'il s'agit de rendre de plus en plus précise l'expression de la pensée. Dans quelques idiomes agglutinants un certain vague est la conséquence de la multiplicité des formes. De là aussi des combinaisons propres à ces idiomes, et qui étonnent singulièrement nos esprits habitués à la simplicité relative des langues indo-européennes. Dans ces dernières langues, les éléments qui ont pour mission d'indiquer la personne, les relations personnelles (*amat*, *il* aime ; *anamus*, *nous* aimons) sont restreints au verbe, à la conjugaison ; de même, les éléments chargés d'indiquer le sujet, l'objet, la place dans l'espace, sont restreints au nom, à la déclinaison : *filius*, sujet ; *filium*, régime direct. Mais le procédé de l'agglutination permettait la formation de dérivés appartenant à d'autres catégories. En magyar, par exemple, le mot *munka* « ouvrage » et le suffixe personnel *m* forment le nom *munkám* « mon ouvrage ». Dans les langues dravidiennes nous retrouvons ce procédé, mais ici, et pour parler des mots de cette espèce, le suffixe personnel apporte pour ainsi dire un sens attributif, une signification d'existence. En tamoul, par exemple, *têvarîr* formé de *têvar* « dieu », pluriel honorifique, et de *îr* terminaison personnelle, signifie : « vous êtes dieu » et peut, du reste, en prenant le sens de « vous qui êtes dieu » se prêter aux procédés qui répondent à ce qu'on appelle la déclinaison. Voici un autre fait bien significatif et bien curieux (encore qu'on ne le retrouve plus que dans les textes anciens). Dans les vieux poëmes tamouls on rencontre des formes telles que *çârndâykku* « à toi qui t'es approché » et qui se décompose ainsi : *çâr* « atteindre », s'approcher, arriver » ; *n* euphonique ; *d*, signe du passé ;

ây « tu, toi », suffixe de la seconde personne ; *k* euphonique et *ku* « à ». Des formations de cette sorte sont tout à fait caractéristiques.

Le toulou, une des langues dravidiennes les moins importantes, offre une particularité que nous ne devons pas négliger de mentionner. Le mot en tamoul, en télinga, en kanara, en mahayâla peut donner naissance à un dérivé causatif par le fait de l'intercalation d'une syllabe particulière entre le radical et l'élément qui indique le temps. En tamoul, par exemple, de *çeyvén* « je ferai » l'on tire *ceyvippên* « je ferai faire ». En toulou le nombre de ces formes secondaires est bien autrement considérable : *malpuve* « je fais », *malpêve*, fréquentatif « je fais habituellement »; *malpâve*, causatif « je fais faire » ; *maltruve*, intensif « je fais vivement ». Par l'intercalation d'une nouvelle syllabe, d'un nouvel élément, chacune de ces formes peut devenir négative : *malpâvuji* « je ne fais pas faire ». Ce phénomène se retrouve dans la langue turque, ainsi que nous le verrons en temps et lieu. Les exemples de ce procédé y sont en nombre considérable, et l'on y dit en un seul mot : « je fais aimer, je puis aimer, je m'aime, ils s'aiment l'un l'autre », et ainsi de suite.

Les langues dravidiennes ne connaissent pas l'article, bien que l'on trouve parfois dans de vieux documents des exemples de pronoms démonstratifs employés avec le sens déterminatif. L'adjectif, toujours invariable, n'est pour l'ordinaire qu'un nom de qualité qui précède constamment le nom auquel il sert d'épithète.

La distinction des genres devait être primitivement inconnue en dravidien. De nos jours, même, elle ne s'applique qu'aux êtres humains qui sont parvenus à l'âge de raison : les noms d'enfants sont neutres dans toutes les langues dravidiennes, et, dans la plupart d'entre elles, également, les noms de femmes le sont aussi au singulier.

Le système dravidien ne conçoit la notion de temps que sous trois hypothèses : celle du présent, celle du passé, celle d'un futur indéterminé et dont l'idée est très-vague; ce futur indique, par exemple ce qui est, était ou doit être fait d'habitude. Quant aux deux « voix » dont parlent les grammairiens, l'une positive, l'autre négative, elles se réduisent à une seule et même forme primitive. La voix négative, en effet, n'est que le composé d'une négation, des éléments chargés d'indiquer la personne et du simple radical.

Le vocabulaire dravidien indique un état de civilisation assez peu avancé. Il n'y avait dans le pays dravidien, avant l'arrivée des populations hindoues, ni « dieu », ni « âme », ni « temple », ni « prêtre ». Il est vrai qu'il n'y avait pas davantage de « livre », d' « écriture », de « grammaire ». Le mot « volonté » fait également défaut. On ne savait pas compter jusqu'à mille ; la seule langue dravidienne qui possède un mot propre pour exprimer ce nombre, le télinga, l'a tiré de la racine *ve* « ardeur, multiplication ». Aucun idiome dravidien ne peut rendre dans leur sens abstrait nos verbes « avoir » et « être ».

D'après cette esquisse, on peut juger suffisamment du caractère des langues dravidiennes. Ce sont des langues agglutinantes arrêtées dans le développement de leurs formes à une période pour ainsi dire prématurée. L'invasion hindoue fut, selon toute vraisemblance, la cause de cet arrêt.

Quoi qu'il en soit, il est aisé d'assigner aux langues qui représentent aujourd'hui le système dravidien leur place naturelle dans la série des idiomes agglutinants. Elles doivent être inscrites parmi les premières par ordre ascendant, c'est-à-dire parmi celles qui suivent immédiatement le monosyllabisme, mais qui précèdent le turc, le magyar, le basque et les langues américaines. L'on n'y

trouve aucune trace de flexion, et les altérations de voyelles qu'il est loisible de constater chez elles sont purement phonétiques. Elles n'ont aucune importance en ce qui concerne le sens même des mots qui sont ainsi modifiés.

Nous avons dit que le contact des langues hindoues avait été la cause probable de l'arrivée des Dravidiens à la vie historique. Tout indique en effet que les Aryens furent à la fois les conquérants des plaines et des forêts du Décan et les civilisations de leurs sauvages habitants. Des tribus errantes et misérables, indisciplinées, difficiles à aborder, peuplent encore quelques contrées à peine explorées de cette riche et féconde région. S'il est à présumer que les Dravidiens ont été civilisés par l'invasion hindoue, il est certain au moins qu'ils lui doivent leur écriture.

Les langues dravidiennes littéraires sont transcrites pour l'ordinaire au moyen de trois alphabets différents. Le toulou emploie les mêmes caractères que le kanara. Ce dernier idiome et le télinga n'ont au fond qu'un seul et même système, et la forme de leurs lettres respectives ne présente que des différences minimes. Le premier des trois alphabets dont nous parlions est cet alphabet kanaro-télinga. Le second est celui du tamoul. La forme carrée y prédomine et il ne possède que vingt-huit signes tandis que les autres reproduisent avec fidélité l'ordre et le nombre des lettres de l'alphabet propre au sanskrit. Aussi les brahmanes du pays tamoul se servent-ils, lorsqu'ils veulent écrire du sanskrit, d'un alphabet spécial appelé « grantha » ; cet alphabet est calqué sur l'écriture hindoue et présente un des deux types anciens d'où est sorti, par voie de réduction, le système graphique tamoul (qui, soit dit en passant, confond la forte et la faible de chaque paire d'explosives). Le troisième alphabet dravidien est celui du malayâla qui dérive du gran-

tha. Les anciennes inscriptions dravidiennes se ramènent à deux types d'écriture : l'une est spéciale au tamoul ; l'autre qui sert au sanskrit et aux langues indigènes et se rapproche considérablement des vieilles formes de l'alphabet du sanskrit, serait le prototype de tous les alphabets du Décan. La première aurait été empruntée directement aux Sémites, selon Burnell.

Des peuples qui ne possèdent point d'écriture peuvent-ils avoir une littérature au sens propre du mot ? En tout cas, il y a maint exemple de populations tout à fait illettrées chez lesquelles de longues compositions, toujours poétiques, se sont transmises de bouche en bouche à travers nombre de générations, et l'on découvre partout des chants et des contes populaires qu'aucune plume, qu'aucun outil n'ont fixés. En est-il de même chez les anciens Dravidiens ? Nous ne pouvons l'affirmer. La littérature des langues dravidiennes est pourtant assez riche ; mais tous les ouvrages dont elle se compose, jusqu'aux moindres fragments, sont postérieurs, et de beaucoup, aux premiers temps de l'influence aryenne. Au point de vue du nombre et de la valeur de ces compositions, le tamoul et le kanara l'emportent sur les autres idiomes dravidiens, bien que le télinga offre aux érudits une curieuse mine à fouiller. La littérature tamoule est cependant la plus abondante, la plus féconde, la plus intéressante et en même temps la plus ancienne. Le tamoul littéraire diffère beaucoup plus du tamoul vulgaire que ne diffèrent les autres idiomes dravidiens littéraires du kanara, du télinga et du toulou employés dans la conversation usuelle. D'ailleurs la littérature tamoule n'est pas toujours un simple reflet de la littérature sanskrite, et elle possède aussi son originalité. Le tamoul a eu la bonne fortune d'être pendant longtemps la langue des sectaires çivaïstes et celle d'hérétiques djâinistes et bouddhistes qui ont beaucoup écrit. Leurs livres sont les chefs-d'œu-

vre de la poésie tamoule ancienne. Il faut ajouter que dans les langues dravidiennes les vieux monuments, ou ceux qui possèdent quelque valeur, sont toujours en vers. La poésie tamoule est plus pure comme langage, plus correcte que la prose, et proscrit avec bien plus de soin l'emploi de mots étrangers ; on constate tout le contraire dans les vers du télinga, du kanara, du malayâla, où abondent les mots tirés des dialectes hindous. Le vocabulaire tamoul, d'ailleurs, est fort riche et possède un grand nombre de synonymes.

La littérature dravidienne est particulièrement originale dans les poëmes moraux, dans les recueils de sentences et d'aphorismes, qui constituent les plus anciens monuments de la poésie tamoule. Elle a produit également de longs poèmes épiques, remarquables par l'exagération et la minutie des détails, et dont la lecture nous semble en général peu attrayante.

Il faut attribuer à une époque plus récente de nombreux chants lyriques pleins d'emphase, des hymnes religieux pleins de monotonie et des récits licencieux. C'est à une époque moins ancienne encore qu'appartiennent des écrits scientifiques presque exclusivement consacrés à l'art médical.

Certains auteurs ont rattaché aux langues dravidiennes le *brahoui*, parlé au nord-ouest de l'Inde, par delà l'Indus, aux environs de Kélat, dans le Béloutchistan. Cette opinion a été contestée. En fait, on ne peut encore se prononcer sur cette question. Cadwell se refuse à voir dans cet idiome une langue dravidienne. En tout cas, le brahoui est surchargé de mots hindous et arabes.

§14. Les langues ouralo-altaïques.

Disons tout d'abord qu'on s'accorde généralement à diviser en cinq groupes principaux les langues ouralo-

altaïques : le groupe *samoyède*, le groupe *finnois*, le groupe *turc* ou tatar, le groupe *mongol*, le groupe *tongouse*.

Nous leur devons une place importante dans cet écrit. Plusieurs d'entre elles ont ou ont eu une valeur littéraire réelle ; toutes sont pleines d'intérêt en ce qui concerne la linguistique proprement dite. On a souvent cité telle ou telle langue de la famille ouralo-altaïque (notamment le turc ou le magyar), lorsqu'il s'agissait d'exposer les procédés de l'agglutination. Elle se prétend, en effet, on ne peut mieux à cette démonstration.

En premier lieu, nous nous proposons de passer en revue les cinq groupes énumérés ci-dessus et les principaux idiomes qui les constituent ; nous traiterons ensuite d'un phénomène d'euphonie qui a une valeur considérable dans cette famille linguistique et auquel on a donné le nom d'harmonie vocalique.

Mais avant tout nous devons noter ce fait important : bien que rangées sous un seul et même titre, les différentes langues ouralo-altaïques offrent entre elles des diversités considérables, non-seulement en ce qui concerne leur vocabulaire, mais encore en ce qui regarde leur structure.

Après avoir parlé du phénomène de l'harmonie vocalique, nous reviendrons sur cette question de la parenté des cinq groupes ouralo-altaïques, mais dès à présent nous n'hésitons pas à les classer, comme on le fait d'habitude, à côté les uns des autres et sous une même rubrique. Tous ces idiomes en effet connaissent plus ou moins les procédés qui consistent à suffixer aux noms un pronom possessif, à joindre un régime direct au mot comportant la notion verbale, et leurs concordances lexiques sont souvent remarquables.

Cela dit, nous entrons de suite en matière.

I. *Le groupe samoyède.*

Il s'étend en Europe sur la partie orientale de la côte russe de l'océan Glacial (c'est-à-dire à l'est de la mer Blanche), et en Asie sur la partie occidentale de la côte sibérienne.

On ne porte pas au nombre de plus de vingt mille les dialectes principaux qui se subdvisent presque tous en sous-dialectes.

Le *yourak* est parlé dans la Russie européenne et dans le nord-ouest de la Sibérie, jusque vers le fleuve Iénisséi.

Le *samoyède iénisséin* occupe la région du bas Iénisséi.

Le *tavghi* est parlé plus à l'est, jusqu'à l'embouchure du Chatanga.

Plus au sud-ouest, et sur le cours moyen de l'Ob, on trouve le *samoyède ostiaque*, vers les rivières Tym et Tchulym.

Enfin le *kamassin* est la langue d'un petit nombre d'habitants de la Sibérie méridionale.

Le Finnois Castrén, l'un des fondateurs de la linguistique ouralo-altaïque, a écrit un travail étendu et méthodique sur les dialectes samoyèdes où il les compare continuellement entre eux (1). Dans sa pensée le samoyède se rapproche du finnois bien plus que d'aucun autre groupe ouralo-altaïque, et cela sous le rapport de la formation des mots comme sous le rapport du matériel même de la langue.

Le système des voyelles est plus compliqué dans les idiomes samoyèdes ; celui des consonnes est au contraire assez développé : on en compte plus d'une trentaine, parmi lesquelles le *t* mouillé, le *d* mouillé, *l* mouillé, *s* et *z* également mouillés.

(1) *Grammatik der samojedischen sprachen.* (Publ. par M. Ant. Schiefner). Pétersbourg, 1854.

Nous parlerons un peu plus loin des principes de l'harmonie vocalique, qui consiste dans l'assimilation de la voyelle des éléments secondaires du mot à la voyelle de la syllabe principale. Ce principe est loin d'être observé également dans tous les dialectes samoyèdes ; ce n'est que dans le dialecte kamassin qu'on le voit bien développé. Dans ce dialecte en effet les voyelles dites fortes (*a*, *u*, *o*) ne peuvent se rencontrer avec les voyelles dites faibles (*ä*, *ü*, *ö*), tandis que les voyelles dites neutres (*i*, *e*) se prêtent parfaitement au voisinage des fortes et des faibles.

Comme dans les autres langues ouralo-altaïques la notion de la déclinaison est rendue en samoyède par l'agglutination à la racine principale des racines secondaires indiquant telle ou telle idée de rapport, de relation. Le suffixe *n*, par exemple, indiquant l'idée du génitif, on dit en samoyède ostiaque *loga* « le renard », *kule* « le corbeau » et *logan* « du renard » *kulen* « du corbeau ». Si au thème vient s'ajouter l'élément *la* indiquant le pluriel, on dit *logola* « les renards », *kulela* « les corbeaux » et *logalan* « des renards », *kulelan* « des corbeaux ». Tout ce mécanisme est des plus simples.

II. *Le groupe finnois.*

Ce groupe est d'un intérêt bien plus considérable que le précédent, et aucun autre ne joue un rôle aussi important dans l'étude des langues ouralo-altaïques. On lui donne le nom d' « ougrien », de « finno-ougrien » ou d' « ougro-finnois ».

Quant aux langues dont il se compose, elles ne sont pas encore distinguées l'une de l'autre d'une façon bien définitive. Donner voit ici deux sous-groupes. La branche

ougrienne comprendrait trois rameaux secondaires : l'ostiaque, le vogoul, le magyar. La branche finnoise comprendrait deux rameaux secondaires : le rameau permien, se subdivisant, d'une part, en votiaque ; de l'autre, en zyriène et en permien, puis le rameau volga-baltique. Ce dernier rameau se subdiviserait lui-aussi : premièrement, en variété du volga, comprenant le tchérémisse et le mordvin ; secondement, en variété du finnois occidental, comprenant tout le reste de la sous-famille, soit le lapon, le live, le vepse, l'esthonien, le vote, et enfin le suomi ou finnois proprement dit (*Die gegenseitige verwandtschaft der finnisch-ugrischen sprachen ;* Helsingfors, 1879).

Le *suomi* occupe la plus grande partie de la Finlande, mais il ne s'étend pas sur toute la côte du golfe de Bothnie qui longe, un peu vers le nord, une bande de territoire où l'on parle suédois, — à Vasa, par exemple ; — au sud, le suomi n'atteint que sur des points peu importants le golfe de Finlande, dont la côte septentrionale — Helsingfors et ses alentours — est également occupée par des Suédois. On trouve encore un certain nombre de Finnois aux environs de Saint-Pétersbourg.

Au suomi l'on rattache le *karélien*, qui s'étend au nord jusqu'au territoire lapon, au sud jusqu'au golfe de Finlande et au lac Ladoga, à l'est jusqu'à la mer Blanche et aux bords du lac Onéga, -- et le *tchoude*, parlé sur un territoire très morcelé, au sud de ce dernier lac : le *vepse* est le tchoude du nord, le *vote* est le tchoude du sud.

L'*esthonien*, ou mieux l'eshte ou este, est beaucoup moins répandu que le suomi et les idiomes qui lui sont apparentés. Il est parlé par huit cent mille individus environ et occupe la plus grande partie de la côte méridionale du golfe de Finlande (Réval, Vésenbourg) ainsi que la moitié septentrionale de la Livonie (Dorpat). On distingue deux dialectes esthoniens, celui de Réval et celui

de Dorpat, se subdivisant à leur tour en plusieurs sous-dialectes, mais ne s'étant jamais prêtés à une langue littéraire commune. On essaya, mais en vain, vers la fin du dix-septième siècle, de constituer un esthonien littéraire ; cette entreprise échoua complètement, comme elle devait échouer (1). La littérature de l'esthonien le cède de beaucoup à celle du suomi.

La langue *live* n'occupe guère plus que la pointe nord-occidentale de la Courlande, quelques lieues carrées à peine. Du côté de la terre elle est de plus en plus pressée par un idiome indo-européen du groupe lithuanien, le lette.

Disons quelques mots ici de la grammaire du suomi (2), puis de la grammaire esthonienne.

Le système des consonnes du suomi est très simple. Outre les explosives *k*, *t*, *p*, il possède les liquides *r*, *l*, les nasales *m*, *n*, et une troisième nasale assez semblable à celle de l'allemand « lang », puis *s*, *h*, *v*, *y*, (que l'on écrit *j*). Il ne connaît ni les explosives aspirées, ni « f ». Les explosives faibles *g*, *d*, *b*, s'y rencontrent, mais on les donne comme étrangères au fond même de la langue ; elles remplaceraient des *k*, *t*, *p* plus anciens.

Le suomi semble aimer l'hiatus, la rencontre de plusieurs voyelles. Toutes les voyelles peuvent finir les mots, sauf, en principe, la voyelle *e*. Il n'en est pas de même de toutes les consonnes. Très souvent c'est la consonne *n* que l'on rencontre à la fin des mots.

Nulle part, le principe de l'harmonie vocalique (dont nous aurons à parler ci-dessous avec plus de détails) n'est plus frappant qu'en suomi. La syllabe radicale du mot contient-elle une voyelle forte, les voyelles des suf-

(1) Wiedeman. *Grammatik der eshtnischen sprache*. Pétersbourg, 1875.

(2) Kellgren. *Die grundzüge der finnischen sprache mit rücksicht auf den ural-altaischen sprachtstamme*. Berlin, 1847.

fixes doivent également être fortes ; contient-elle une voyelle faible, les voyelles des suffixes doivent être faibles ; contient-elle une voyelle neutre, moyenne, les voyelles des suffixes doivent être faibles.

Notons qu'en suomi la formation des mots n'a jamais lieu par des préfixes, c'est-à-dire que la racine occupe toujours la première place et que les éléments qui la dérivent sont agglutinés à la suite de cette même racine, et non devant elle. D'autre part, de même qu'en magyar, c'est sur la racine, c'est-à-dire sur la première syllabe du mot, qu'est placé l'accent principal.

En somme le suomi est une langue très euphonique. Il assimile volontiers les consonnes, notamment celles qui terminent la racine et celles qui commencent les éléments dérivatifs, les suffixes. A vrai dire, cette assimilation n'est pas constante, mais lorsqu'il l'évite il recourt, pour échapper au heurt de deux consonnes d'ordre différent, à un autre procédé. Ce dernier consiste à introduire dans la prononciation (sinon l'écriture) une voyelle très brève entre les deux consonnes en question. Ainsi *pitkä* « long » se prononce *pitikä*.

Les relations diverses que les langues à flexion expriment en principe par leur cas sont rendues en suomi, comme dans les autres idiomes altaïques, par l'agglutination de différents suffixes à la forme radicale du mot. Pour exprimer par exemple l'idée du génitif, on emploie le suffixe *n* : *karhu* « l'ours », *karhun* « de l'ours ». Le suffixe chargé d'exprimer le pluriel est *t* lorsque le mot est sujet de la phrase ; dans les autres circonstances, c'est *i* qui se place entre le radical et le suffixe indiquant la relation. Ainsi le thème *lapse* « enfant » donna naissance aux formes suivantes : *lapsen* « de l'enfant », *lapset* « les enfants », *lapsein* « des enfants ».

Le suomi annexe les pronoms personnels au nom substantif lorsqu'il s'agit de dire à quelle personne se

rapporte ce nom. Pour la première personne le pronom ainsi suffixé est *ni* au singulier, *mme* au pluriel ; pour la seconde personne c'est *si* au singulier, *nne* au pluriel ; pour la troisième c'est *nsa* (ou *nsä* d'après les principes d'euphonie) pour le pluriel comme pour le singulier. C'est ainsi que de *tapa* « coutume » on forme *tapani* « ma coutume », *tapamme* « nos coutumes », *tapansa* « sa coutume » ou « leurs coutumes ».

C'est également par une série de suffixes que le suomi forme ce que l'on nomme son verbe. La racine toujours invariable se place au commencement du mot, puis viennent les suffixes indiquant que ce mot est causatif, diminutif, fréquentatif, puis les suffixes indiquant le mode, puis les suffixes indiquant la personne, le sujet de l'action.

Le système des consonnes de l'esthonien n'offre rien de particulier si ce n'est que les *t*, *d*, *n*, *r*, *l*, *s*, *z*, sont mouillés en certaines circonstances. C'est ce que l'on indique dans l'écriture au moyen d'un signe-minute adjoint au caractère : *d'*, *n'*, et ainsi de suite. Le dialecte de Dorpat prononce les *g*, les *d*, les *b* plus énergiquement que ne le fait l'autre dialecte ; il les change parfois en leurs correspondantes fortes, *k*, *t*, *p*. Parmi les neuf voyelles esthoniennes on trouve *ü* (« u » français) et un son spécial assez rapproché de *ö* et de *e*. Tantôt ces voyelles sont longues, tantôt elles sont brèves ; souvent elles forment diphtongues.

En parlant des dialectes samoyèdes et du suomi, nous avons dit quelques mots de l'harmonie vocalique, à laquelle nous consacrerons d'ailleurs, ci-dessous, un paragraphe spécial. Le phénomène de l'harmonie vocalique, de l'analogie qui doit se rencontrer entre la voyelle des éléments dérivatifs et la voyelle de la syllabe radicale, est loin d'être généralisé en esthonien. Il ne se présente réellement que dans la partie orientale du dialecte de

Dorpat, mais on en trouve cependant des traces évidentes dans la partie occidentale du même dialecte ainsi qu'à l'ouest et au sud du dialecte de Réval.

L'accent principal en esthonien tombe sur la première syllabe ; cette syllabe est en réalité la syllabe radicale.

Le procédé de formation des mots auquel on donne abusivement le nom de déclinaison se prête ici, comme dans les autres langues agglutinantes, à une foule de soi-disant cas. Le nombre de ces derniers n'est limité que par le nombre même des postpositions que l'on peut joindre au mot. En esthonien on en compte tantôt une dizaine, tantôt une vingtaine. Cette hésitation des grammairiens suffit à démontrer combien ces prétendus cas diffèrent des cas réels de la déclinaison des langues indo-européennes ou de celle de l'arabe littéral.

De ce que l'on appelle la conjugaison nous ne dirons d'autre part qu'une seule chose, c'est qu'elle est tout à fait analogue à celle du suomi.

Le *lapon* occupe l'extrême nord-ouest de la Russie (au nord du karélien) et quelques régions du nord de la Suède et de la Norvège. On y reconnaît quatre dialectes. Nous ne dirons rien de particulier de sa grammaire qui concorde d'une façon très nette avec celle du suomi et celle de l'esthonien, dont nous venons de nous occuper.

Les idiomes finnois du Volga se divisent en deux branches : le *tchérémisse* et le *mordvin*.

Le tchérémisse est parlé par deux cent mille individus environ, sur la rive gauche du Volga. Le territoire qu'il occupe est assez rapproché de Kazan à l'ouest, de Nijni Novgorod à l'est, mais cependant il n'est point contigu à ces deux villes. On reconnaît dans le tchérémisse deux dialectes : un dialecte de la plaine, un dialecte de la montagne.

Le mordvin est parlé dans un certain nombre d'îlots peu considérables, par près de sept cent mille individus.

On le rencontre à l'est et à l'ouest du Volga, à la hauteur de Simbirsk, de Stavropol, de Samara et même un peu plus au sud. Il se divise en deux dialectes, l'erza et le mokcha.

Entre le mordvin et le tchérémisse, se trouve placé le tchouvache, idiome ouralo-altaïque lui aussi, mais qui appartient au groupe turc ou tatar, non point au groupe finnois qui nous occupe actuellement.

Le *permien*, parlé par environ soixante mille individus, le *zyriène* par quatre-vingt mille (ou peut-être davantage), le *votiaque* par plus de deux cent mille, se rencontrent plus au nord. Le votiaque occupe un territoire relativement assez compacte, au nord-est du tchérémisse, au sud de Glasov. Le permien s'étend au nord du votiaque, à l'ouest de la rivière Kama, à la hauteur de Solikamsk. Le zyriène, plus au nord que ses deux congénères, occupe un territoire beaucoup plus vaste, et, confinant du côté est au vogoul dont nous parlerons tout à l'heure, il atteint au nord la limite du samoyède.

On donne le nom de rameau « ougrien » au vogoul, à l'ostiaque et au magyar.

Le *vogoul* est parlé par environ sept mille individus, l'*ostiaque* par une vingtaine de mille. Le premier s'étend à l'est du zyriène, sur des régions très peu peuplées ; le second, plus à l'est encore, occupe sur une grande longueur les rive de l'Ob et confine à la limite méridionale du samoyède. Le vogoul comprend au moins deux dialectes ; quant à l'ostiaque, il varie à Irkutsk, à Surgut, à Obdorsk.

Nous devons nous arrêter sur le *magyar* avec plus d'attention. La position géographique, les relations politiques des cinq millions d'individus qui parlent cette langue, sa littérature, assez remarquable, lui donnent une place spéciale parmi tous les autres idiomes du groupe finnois.

Le magyar ou hongrois s'étend sur deux régions d'inégale grandeur et séparées l'une de l'autre par un espace d'environ quarante-cinq à cinquante lieues.

Le groupe principal, celui de l'ouest, présente la forme d'un pentagone irrégulier aux angles duquel se trouvent placées les villes de Presbourg (en magyar Pozsony, où confinent l'allemand à l'ouest et au sud, le slovaque au nord, le magyar à l'est) ; Unghvar, de langue slovaque ; Nagy Banya, de langue magyare ; Novi Sad (en allemand Neusatz, sur la limite du magyar et du serbe de Sirmie) ; Unt. Limbach, de langue magyare (un peu au nord de Varasdin, en Croatie). La plus grande hauteur de ce pentagone est d'environ quatre-vingts lieues, sa plus grande largeur de cent et quelques lieues. Il ne forme pas d'ailleurs un territoire compact ; il renferme un certain nombre d'enclaves de langue slovaque au nord, de langue serbe au sud, de langue allemande à l'ouest et au sud.

Le groupe magyar oriental est plus homogène ; six fois moins étendu, approximativement, que le groupe de l'ouest, il se trouve (avec deux îlots de langue allemande accolés à sa frontière occidentale, Mediasch, Kronstadt) situé juste au milieu de la région de langue roumaine. Ce second territoire magyar, dont les principales localités sont peu considérables (Maros-Vasarhely, Udvarhely), forme l'extrême sud-est du royaume de Hongrie.

On a cherché à expliquer de diverses façons, mais toujours sans succès, le mot de « magyar ». Celui de « hongrois » nous paraît tout aussi obscur. Rappelle-t-il l'origine des Magyars, c'est ce que nous ne pouvons décider. La question, d'ailleurs, n'a qu'une importance secondaire (1).

(1) Sayous. *Les origines de l'époque païenne de l'histoire des*

On suppose, avec assez de vraisemblance, que l'invasion d'Attila ne fut qu'une première incursion des peuples proches parents des Magyars actuels. Quoi qu'il en soit, ceux-ci sont absolument isolés aujourd'hui des autres populations de langue finnoise et se trouvent enveloppés de tous côtés par l'allemand, le roumain et différents idiomes slaves ; il est hors de doute que dans un petit nombre de siècles leur propre langue aura vécu, en dépit des privilèges que les conditions politiques lui auront octroyés à profusion. Elle ne disparaîtra pas d'ailleurs sans laisser une histoire honorable. Son monument le plus ancien est de la fin du douzième siècle. Entre autres écrits du quinzième siècle, on possède une version de la Bible, et parmi ceux du seizième, une légende de sainte Marguerite. Nous n'avons pas à nous étendre ici sur l'histoire littéraire du magyar ; les productions de cette langue restent malheureusement lettre close pour la plupart des érudits et des lettrés étrangers. Cela est particulièrement regrettable à notre point de vue. Nombre de bons écrits grammaticaux sur les langues ouralo-altaïques sont rédigés en magyar et sont condamnés par là même à ne se répandre que très lentement. Les savants hongrois sont habitués, aujourd'hui encore, à manier assez bien la langue latine ; que ne l'emploient-ils pour nous faire connaître leurs propres travaux ?

On compte un certain nombre de dialectes magyars, les uns appartenant à la basse Hongrie (dialectes de Debreczin, de Szegedin, etc.) et ceux de la haute Hongrie. Toutefois, leurs différences sont relativement minimes, et l'on peut dire que la langue magyare n'a pas varié d'une façon considérable depuis l'époque de ses plus anciens documents historiques. Elle a été influencée tour

Hongrois. Paris, 1847. Riedl. *Magyarische grammatik*. Vienne, 1858. Introduction. Castrén. *Ueber die ursitze des finnischen volkes*. Helsingfors, 1849.

à tour, il est vrai, par le turc, par les idiomes slaves, par l'allemand, par d'autres langues encore, mais son fond est resté essentiellement le même.

Depuis que le royaume de Hongrie a repris une importance toute nouvelle, aux dépens des provinces serbes et roumaines de la monarchie austro-hongroise, le magyar a gagné également une nouvelle importance, au moins à ce point de vue. Mais ses productions ont été trop souvent pensées en allemand, et elles se ressentent profondément de l'éducation étrangère de leurs auteurs.

Jetons sur la phonétique et la structure de cette langue un rapide coup d'œil.

Le matériel phonique du magyar est peu compliqué : sept voyelles brèves, *a*, *e* (plus ou moins ouvert), *i*, *o*, *u*, (« ou » français), *ö* (« eu » de « feu »), *ü* (« u » français, « ü » allemand) et leurs sept correspondantes longues ; ces dernières sont marquées dans l'écriture par la superposition d'un accent incliné à droite : *á*, *é*, *ó*, *ú*, etc. Les consonnes ne sont pas fort nombreuses, mais quelques-unes d'entre elles, *ty* et *gy*, n'ont point de correspondante française. Il serait plus simple, sans doute, de les rendre par un signe unique (par exemple *c'* et *j'*). A coup sûr la transcription est détestable, lorsqu'il s'agit des *sz*, *zs* et *s* magyar. En effet, *sz* a le son de notre « s » dans « sœur, sa, son » ; *zs* celui du *z* croate et tchèque, c'est-à-dire « j » français ; *s* celui du *s* croate et tchèque, soit le « ch » français de « chercher ». Il serait malheureusement inutile, à l'heure actuelle, de tenter la réforme de ce système défectueux. Les Slaves ont compris, depuis longtemps, l'importance d'une modification de cette espèce et l'ont réalisée en très grande partie (1) ; mais il y

(1) Picot. *Tableau phonétique des principales langues usuelles*. Revue de linguistique, t. VI, p. 363. Paris, 1874. Dans ce tableau, le magyar et le turc représentent les langues ouralo-altaïques.

a peu d'espoir que les magyars accèdent à une réforme semblable.

De même qu'en suomi, la racine, dans le mot magyar, occupe la première place ; il est rare qu'elle se trouve précédée d'un préfixe. Le mot magyar, en principe, est donc formé sur ce type dont nous avons parlé ci-dessus : R+R ou R +R+R, etc., c'est-à-dire racine+suffixe, ou racine+suffixe+suffixe, et ainsi de suite. La forme préfixe+racine ou préfixe+racine+suffixe est relativement rare. Elle serait due à l'influence des langues indo-européennes, et l'histoire elle-même de l'idiome magyar enseigne qu'elle est récente.

Le magyar est soumis à des lois d'harmonie vocalique. C'est là un sujet particulier sur lequel nous ne devons pas nous arrêter ici et que nous examinerons d'ensemble après avoir parlé de toutes les langues ouralo-altaïques. Rappelons simplement que l'harmonie vocalique consiste en ce fait, que les voyelles des éléments suffixés à la racine s'assimilent à la voyelle de cette racine.

Tout comme en finnois, l'accent se pose en magyar sur la syllabe radicale, qui est placée, ainsi que nous l'avons dit, en tête du mot. La racine est-elle précédée d'un préfixe (nous avons dit également que cette hypothèse était fort rare), c'est sur ce préfixe que tombe l'accent. En d'autres termes, l'accent pèse sur la syllabe initiale, sur la première syllabe du mot.

La formation des mots, c'est-à-dire la dérivation de la racine, est des plus simples. Le pluriel est indiqué par un élément agglutiné : *ház-hak* « les maisons », *atyá-k* « les pères ». Quant aux éléments indiquant la notion de cas, ils ne viennent se juxtaposer qu'après ce signe du pluriel. Au singulier, par exemple, l'on dit *atya* « pater » et *atyat* « patrem », au pluriel *atyák* et *atyákat*, le premier rendant le nominatif latin, le second rendant l'accusatif.

Un article s'est formé dans la langue magyare. C'est *az*

devant les voyelles, *a* devant les consonnes : *az ember* « l'homme ».

La dérivation verbale du magyar est assez riche : c'est une suite d'éléments agrégés les uns aux autres. Elle incorpore (ainsi que cela se passe dans toutes les langues du groupe finnois) le pronom de la troisième personne lorsqu'il est régime direct :

vár « il attend », forme simple au singulier ;

várja « il l'attend », forme objective au singulier ;

várják « ils l'attendent », forme objective au pluriel.

Dans ces différents exemples, l'élément *ja* indique le pronom régime « lui », et *k* est le signe du pluriel.

Et non seulement le magyar peut incorporer le pronom régime de la troisième personne, mais il peut incorporer également celui de la deuxième. Cela n'a lieu, il est vrai, que dans l'hypothèse où le sujet est à la première personne du singulier. A côté de *várok* « j'attends », il dit *várlak* « je t'attends ». Cette faculté d'incorporer le régime de la seconde personne est un fait à relever soigneusement. Nous verrons plus loin que le basque fait encore mieux sous ce rapport que les langues ouralo-altaïques et qu'il incorpore non seulement le régime direct (comme dans « je le donne »), mais encore le régime indirect (comme dans « je te le donne »).

Tous les idiomes que nous venons de passer rapidement en revue, sont incontestablement parents les uns des autres et se rattachent à une souche commune. On a commencé à appliquer à leur étude la méthode comparative, mais ce travail est des plus délicats. Il s'agit ici d'idiomes séparés les uns des autres depuis de longs siècles et qui ont subi l'influence presque continuelle des langues indo-européennes, dont l'organisation est supérieure à la leur propre.

La comparaison des diverses langues finnoises révèle des variations phonétiques curieuses, mais elle ne mon-

tre, en somme, rien de bien nouveau. Citons quelques exemples. Le mot « main » se dit *käte* en suomi, *käzi* en vepse, *tchäsi* en vote, *käsi* en esthonien, *käiz* en live, *giet* et *kät* en lapon, *ki* en zyriène, en permien et en votiaque, *ked* en mordvin, *ket* en tchérémisse, *kèt* ou *kôt* en ostiaque, *kat* en vogoul ; « poisson » est *kala* en suomi, *guolle* en lapon, *kal* en mordvin, *kul* en vogoul, *hal* en magyar.

En général, il semble que le magyar a réduit et abrégé les mots primitifs, tandis que le suomi montre, au contraire, une tendance très accusée à multiplier les voyelles. La comparaison de l'ancien et du nouveau magyar révèle des faits analogues et nous montre que cette langue a subi dans son propre sein des permutations qui sont aujourd'hui normales entre elle-même et ses congénères.

F. Budenz compte dans l'ensemble des langues ougriennes neuf voyelles et trente-trois consonnes, parmi lesquelles le « j » français, les deux « ch » allemands, les semi-linguales de l'ostiaque, le *d* faible et sifflant du lapon et les consonnes mouillées. Aucun des dialectes de cette famille n'a de système graphique véritablement original.

Les idiomes finnois ne connaissent réellement pas la distinction des genres, mais ils possèdent les trois nombres. Le duel et le pluriel sont indiqués par des suffixes différents.

Ainsi que nous l'avons dit, l'article n'est employé d'une façon conforme à nos habitudes qu'en magyar : *az* devant les voyelles, *a* devant les consonnes. Le mordvin, toutefois, sait déterminer les noms, comme le fait la langue basque, en leur postposant le prénom démonstratif de la troisième personne. Le zyriène et le votiaque ont quelque chose d'analogue, et F. Budenz retrouve également des traces de ce procédé dans d'autres langues de

la même famille. En magyar, en effet, l'affixe de la troisième personne, *a* ou *e* « son », est un élément dérivatif très commun : *Pest városa* « la ville de Pesth », mot à mot « Pesth sa ville ».

Il est bon de le répéter, de même que dans toutes les langues agglutinantes, il n'y a point ici de véritable déclinaison. On se sert de postpositions, de particules, dont le sens est celui de nos prépositions, et que l'on place à la fin des mots ; dans l'écriture on a pris l'habitude de ne point les en séparer. Les augmentatifs, les diminutifs, les superlatifs sont formés de la même façon. Les suffixes qui correspondent aux cas de nos langues indo-européennes anciennes, se placent toujours les derniers, par la raison bien simple qu'ils n'affectent pas le sens intime du mot, mais qu'ils indiquent seulement sa manière d'être (à, dans, avec, de) vis-à-vis des autres termes de la proposition. Le nombre de ces suffixes est considérable, aussi les auteurs qui s'avisent de rédiger des grammaires de langues agglutinantes à la façon des grammaires du grec et du latin, ont-ils imaginé de leur donner une quantité de prétendus cas, pour lesquels ils ont inventé les noms en « if » les plus baroques. Il eût mieux valu parler simplement de suffixes ou de postpositions.

Nous avons traité également de la remarquable faculté d'incorporer dans le nom le pronom qui lui sert de régime direct : « je le vois, je le prie », et nous avons dit qu'en un certain cas le magyar pouvait incorporer le pronom de la seconde personne, « je te vois, je te prie ». Le vogoul fait de même, quel que soit le sujet : « je te vois, il te voit » ; le mordvin, enfin, incorpore même le régime de la première personne et dit en un seul mot « il me voit, il me prie ». Notons d'abord que ces idiomes possèdent également les formes où ce régime n'est pas incorporé.

Ces observations sur le groupe finnois des langues ou-

ralo-altaïques sont assez succinctes, mais elles suffisent, nous semble-t-il, à le caractériser et à montrer quelle est son importance, aussi bien que son intérêt.

III. *Le groupe turc.*

On lui donne aussi le nom « tatar », et fort improprement celui de « tartare », qui n'est qu'un mauvais jeu de mots. Les peuples qui parlent les nombreux idiomes formant le groupe turc s'étendent aujourd'hui sur les rivages de la Méditerranée orientale aux bords de la Léna en Sibérie. On enseigne communément que leur point d'origine fut le Turkestan (1) ; c'est de là que rayonnèrent, depuis les âges historiques, des hordes nombreuses et intrépides qui conquirent en Asie de vastes régions et poussèrent en Europe jusque sur le territoire français.

Au point de vue linguistique, les Turcs, dans la plus large extension du mot, se partagent en cinq familles, parlant chacun un idiome distinct divisé à son tour en un plus ou moins grand nombre de variétés. En allant de l'est à l'ouest et du nord au sud, ces cinq rameaux sont le *yakout*, *l'ouigour*, le *nogaïque*, le *kirghiz*, le *turc* proprement dit.

Le *yakout* est parlé par deux cent mille individus environ, au milieu des peuplades tongouses dans la Sibérie du nord-est.

On compte trois dialectes ouigours : l'*ouigour* proprement dit, le *djagataïque*, le *turcoman*. La langue ouigoure est de toutes ses congénères celle qui a atteint le

(1) Abel Rémusat. *Recherches sur les langues tatares*, p. 328. Paris, 1820.

plus haut degré de culture littéraire. Elle s'écrivait encore au cinquième siècle de notre ère, comme en témoignent les auteurs chinois, à l'aide d'un alphabet original, perdu depuis lors et remplacé, sous l'influence des missionnaires nestoriens, par un système dérivé de l'alphabet syriaque, comme celui des Mandchous, des Kalmouks, des Mongols (1).

Le *nogaïque* est parlé par environ cinquante mille personnes vers l'embouchure du Volga, à Astrakan, dans quelques districts situés entre la mer Noire et la mer Caspienne, dans un petit territoire au nord de la mer d'Azov et dans toute la Crimée. C'est la langue des Tatars proprement dits. Le dialecte koumuque est parlé au nord-est du Caucase.

Certains auteurs rattachent le *kirghiz* au nogaïque. Les Kirghizes noirs, ou Bouroutes, habitent la partie du Turkestan qui se rattache à la Chine. Les Kirghizes kasaks s'étendent plus à l'ouest jusqu'au lac d'Aral et jusqu'au nord de la mer Caspienne.

La cinquième famille est celle des dialectes turcs proprement dits. On y rattache le *tchouvache*, parlé, comme nous l'avons dit plus haut, entre deux idiomes finnois, le mordvin et le tchérémisse. Il occupe un territoire compact assez important au sud-ouest de Kazan et un grand nombre de petits îlots disséminés aux environs de Simbirsk. Le tchouvache offre des particularités remarquables, et certains auteurs l'ont regardé comme un mélange de turc et de finnois, ce qui, d'ailleurs, n'est pas exact. W. Schott a démontré clairement qu'il appartenait au groupe turc, mais pour certains auteurs il se rattacherait, dans ce groupe, non pas au turc proprement dit, mais bien au nogaïque.

(1) Abel Rémusat. *Op. cit.*, p. 254.

Le turc, qui, de tous les dialectes de ce groupe, est le plus intéressant pour les Européens, ne doit pas être considéré comme le plus pur et le plus correct des idiomes de sa famille. Il varie d'ailleurs d'une façon très marquée dans les différentes localités où il est parlé ; la langue des hommes du peuple de Constantinople est beaucoup moins mélangée d'éléments empruntés à l'arabe que ne l'est celle du lettré, du fonctionnaire, de l'*osmanli*. C'est de cette dernière langue que nous allons donner une rapide esquisse. On peut d'ailleurs considérer l'osmanli comme le type le plus frappant d'un idiome agglutinatif, tant sa structure est claire et précise. Les grammaires turques ne font pas défaut, mais la plupart d'entre elles sont faites sans critique ; nous nous servons spécialement de celle de Redhouse (1).

Le turc s'écrit à l'aide de l'alphabet arabe, qui lui convient pourtant aussi peu que possible. Nous avons dit plus haut, et nous le montrerons tout à l'heure, que les voyelles jouent un rôle des plus considérables dans les langues ouralo-altaïques ; or, l'écriture arabe se prête fort mal à la distinction des voyelles. L'alphabet turc se compose de trente et un caractères, susceptibles de recevoir chacun douze signes modificatifs, dont les uns représentent les diverses voyelles, tandis que les autres indiquent que la consonne qu'ils accompagnent doit être prononcée double, ou doit, au contraire, n'être pas prononcée du tout. Sans nous arrêter à ces considérations, si nous recherchons quels sont les éléments phonétiques de la langue turque, nous y trouvons sept voyelles simples : *a*, *e*, *ô*, *u* (prononcez « ou »), *eu* (en un seul son), deux *ü* (« u » français), l'un bref, l'autre long ; notre voyelle nasale « in » (de « maintien ») ; la demi-voyelle

(1) *Grammaire raisonnée de la langue ottomane*. Paris, 1846.

y. Nous y trouvons aussi vingt-deux consonnes, dont un certain nombre de soufflantes gutturales et de sifflantes. Redhouse envisage le vocalisme turc un peu différemment ; on consultera avec fruit, à ce sujet, les tableaux de E. Picot dont nous avons parlé ci-dessus.

La langue turque est toute subordonnée à une règle impérieuse d'harmonie vocalique, sur laquelle nous reviendrons dans un des prochains paragraphes, et qui s'applique même, chez elle, aux mots empruntés à l'arabe et au persan.

C'est en vertu de cette règle que la finale des infinitifs est en *maq*, si la voyelle accentuée du mot est forte ; en *mek*, si cette voyelle est faible ; on dit, par exemple, *sevmek* « aimer » et *yazmaq* « écrire ».

La distinction des genres que l'on reconnaît en turc pour les mots d'origine persane ou d'origine arabe est tout étrangère, primitivement, aux langues tatares. Le turc n'a que deux nombres : le singulier et le pluriel, mais il conserve leur duel aux mots empruntés par lui à l'arabe. Comme toutes les langues ouralo-altaïques, il exprime ce que nous appelons les cas dans les langues indo-européennes, par des postpositions, c'est-à-dire par des syllabes indépendantes placées à la fin des mots et jointes à ces mots dans l'écriture.

Le signe du pluriel, qui est *lar* ou *ler* suivant la nature de la voyelle dominante du mot, s'intercale entre le nom et les suffixes qui sont postposés : *oda* « la chambre », *odada* « dans la chambre » ; pluriel : *odalar* « les chambres », *odalarda* « dans les chambres ». C'est le procédé général des langues ouralo-altaïques.

La véritable nature de ces suffixes terminaux est si bien celle de nos prépositions, qu'un seul d'entre eux suffit pour une série de mots subordonnés, tels, par exemple qu'un substantif joint à des adjectifs.

Quelques-uns de ces suffixes ont d'ailleurs une exis-

tence propre, une existence indépendante, et servent comme des noms communs dans le langage habituel.

L'adjectif, qui n'est qu'un nom qualificatif, se place toujours avant le nom auquel il sert d'épithète. Les degrés de comparaison sont exprimés par l'adjonction de mots au sens de « plus, davantage, moins », etc.

Quant aux prénoms, ils sont ou bien isolés, ou bien joints au nom : *tefter* « cahier », *tefterim* « mon cahier ». En ajoutant l'élément du pluriel, nous avons *tefterlerim* « mes cahiers », en ajoutant la postposition locative *tefterimde* « dans mon cahier ». On voit que ce procédé est des plus simples.

On a souvent cité pour la richesse et la variété de ses formes ce que l'on appelle la conjugaison du turc. Pourtant, il faut le reconnaître, malgré le vaste échafaudage de temps, de modes et de voix dérivées qu'édifient les grammairiens, le finnois ici l'emporte nettement sur le turc. Le magyar, par exemple, formule en un seul mot cette proposition « il l'attend » : *várja* (comparer *vár* « il atend »), en incorporant le régime direct. Le turc ne saurait en faire autant.

Sa grande originalité, c'est ce que l'on appelle ses voix dérivées, c'est la façon dont il combine les formes exprimant diverses nuances des manières d'être d'une seule et même action. A la racine pure et simple, il ajoute, dans ce but, un certain nombre de suffixes (dont la voyelle varie, bien entendu, selon les règles de l'harmonie). Voici quelques exemples de cette faculté de combinaison.

La forme *sevmek*, ainsi que nous l'avons dit tout à l'heure, signifie « aimer » : étant donnés les suffixes *ma*, *me*, indiquant la négation ; *dir*, la causalité ; *il*, le passif ; *in*, le sens réflectif, nous trouvons les formes *sevmemek* « ne pas aimer », *sevdirmek* « faire aimer », *sevilmek* « être aimé », *sevinmek*, « s'aimer », *sevinmemek*, « ne pas s'aimer », *sevdirmemek*, « ne pas faire aimer ».

Chaque racine pourrait fournir de cette façon une cinquantaine de formes dérivées.

C'est également par l'intercalation de certains éléments entre la racine et la terminaison personnelle, que l'on arrive à formuler les notions de temps ou de modalité. Mais à côté de ce procédé très naturel, le turc en possède un autre non moins simple, fondé sur la substitution d'une périphrase à une forme simple, et qui consiste à unir les divers participes avec l'auxiliaire « être ». On peut exprimer de la sorte une foule de nuances très variées et très fines.

Nous ne parlerons pas de la syntaxe turque ; cela nous conduirait hors des bornes de ce travail. Nous pouvons dire toutefois qu'elle est d'autant plus compliquée, que l'idiome s'est notablement altéré par l'intrusion de mots étrangers. Il en résulte que les grammaires sont pleines de règles dont les unes ne s'appliquent qu'aux mots persans, dont les autres ne s'appliquent qu'aux mots arabes et dont quelques-unes sont communes à ces deux catégories d'éléments, sans s'appliquer davantage aux mots d'origine tatare. Ajoutons que le vocabulaire ottoman est profondément mélangé de mots sémitiques et indo-européens empruntés successivement aux Persans et aux Arabes.

Le turc est parlé, en Asie, dans l'intérieur de l'Asie Mineure : la côte appartient à la langue grecque, aussi bien au nord et au sud que sur la mer de Marmara. En Europe, il ne s'étend que sur une faible partie de l'empire ottoman, où il n'occupe nulle part, d'ailleurs, des contrées bien considérables. Ses îlots les plus importants sont situés au sud et à l'est de l'empire : en pays de langue grecque, à Larisse en Thessalie, ça et là dans la Thrace ; en pays de langue bulgare, dans quelques îlots disséminés autour de Philippopoli, et spécialement au nord-est de la péninsule, en dessous de Silistrie. Dans l'île de

Candie, le turc possède encore, au centre même de l'île, un petit territoire assez compacte ; mais, là aussi, la langue grecque gagne sur lui.

IV. *Le groupe tongouse.*

Ce groupe comprend trois branches distinctes : le *mandchou*, le *lamoute*, le *tongouse* proprement dit.

Les Tongouses, au nombre d'environ soixante-dix mille individus, habitent dans la Sibérie centrale. Les Lamoutes, qui se rattachent aux Mandchus, s'étendent plus au nord-est. Les Mandchous occupent le nord-est de la Chine.

Le système d'écriture du mandchou est assez curieux. Il dérive de l'écriture syriaque. Son alphabet comprend vingt-neuf signes ayant chacun une forme triple, comme cela se présente dans l'arabe : on y distingue, en effet, les lettres initiales, les lettres médianes et les lettres terminales. D'ailleurs, ces trois sortes de caractères diffèrent peu les uns des autres. Il s'y ajoute quelques signes complexes dérivés du chinois et destinés, vraisemblablement, à la transcription des mots d'emprunt. Les caractères mandchous sont formés, pour la plupart, d'une barre et d'appendices recourbés ; ils s'écrivent de haut en bas, et les lignes se suivent de gauche à droite. On reconnaît là l'influence des Chinois. Quant aux dialectes tongouses de la Sibérie, ils ne possèdent pas de système graphique particulier.

Il y a peu à dire sur les voyelles du mandchou, mais son système de consonnes est assez compliqué et la classification n'en est pas facile. On y trouve de doubles *k*, *g*, *h*, *t*, *d*, dont les uns ne peuvent se joindre qu'aux voyelles fortes « a, o, ô », les seconds qu'aux voyelles dites neutres « i, u » et à la voyelle dite faible « e » (distinc-

tion dont nous parlerons tout à l'heure en nous occupant de l'harmonie vocalique). Un phénomène assez curieux (et qui se retrouve dans les langues dravidiennes), c'est que les mots ne peuvent commencer par les explosives faibles « b, d, g ». Le mandchou connaît les sons « tch, dj » du français, plusieurs « n », différentes sifflantes (1). Les dialectes sibériens, chez lesquels les consonnes douces peuvent se présenter au commencement des mots, emploient un bien plus grand nombre de sons ; on y trouve quantité de consonnes mouillées, analogues, par exemple, aux « gy, ty, ly » du magyar. L'accent tombe sur la dernière syllabe.

En mandchou, le nom n'a ni genre ni nombre ; les dialectes tongouses, mieux avisés, ont conservé un signe du pluriel : *bira* « la rivière », *biradu*, « dans la rivière », *biral*, « les rivières », *biraldu*, « dans les rivières ». En mandchou, c'est par un moyen syntaxique que s'exprime l'idée du pluriel ou par l'emploi simultané d'un mot indiquant cette notion.

Comme dans les autres idiomes altaïques, c'est par l'adjonction de suffixes à la racine principale que se rend l'idée des cas, l'idée des prépositions du français : mandchou et tongouse *bira* « la rivière », mandchou *birade*, « dans la rivière », tongouse *biradu*, même sens.

La dérivation à idée verbale est analogue à celle du turc et des autres langues agglutinantes ; on y rencontre une quantité de formations secondaires. La racine « boire » donne, par exemple, des formes dérivées dont le sens est « faire boire, venir de boire, aller boire, boire ensemble », et ainsi de suite. Dans toute cette importante partie de la grammaire, les dialectes sibériens procèdent exactement comme le mandchou, mais avec plus

(1) L. Adam. *Grammaire de la langue mandchou*. Paris 1872. Du même auteur : *Grammaire de la langue tongouse*. Paris, 1874.

d'abondance. Ils ont notamment un plus grand nombre de voix dérivées.

Le vocabulaire mandchou-tongouse est assez pauvre, ainsi qu'il est facile de le supposer. Nous voyons qu'il ne connaît pas à proprement parler de verbe « avoir », et nous constatons qu'il a emprunté beaucoup de mots au chinois en les altérant plus ou moins.

La question de savoir qui peut réclamer la priorité sur son congénère, soit du mandchou, soit du tongouse proprement dit, est tranchée par Lucien Adam au profit de ce dernier idiome. Le tongouse possède en effet le signe de pluralité, les pronoms possessifs affixes et d'autres éléments importants qui sont étrangers au mandchou. Les deux idiomes d'ailleurs sont très proches parents, comme le démontre l'identité constante des principaux pronoms, des noms de nombre, des suffixes les plus importants et de la grande généralité du vocabulaire. Evidemment, ils sont issus d'une même souche et leur séparation n'a eu lieu qu'après une assez longue période de développement grammatical commun.

V. *Le groupe mongol.*

Il comprend trois dialectes. Le *mongol oriental*, parlé dans la Mongolie proprement dite, c'est-à-dire dans la partie centrale du nord de la Chine, à l'ouest du territoire mandchou. Le *kalmouk*, ou mongol occidental, qui a pénétré en Russie jusque sur la rive gauche de la mer Caspienne, vers l'embouchure du Volga, entre le kirghiz et le nogaïque, appartenant tous deux, ainsi que nous l'avons dit, au groupe turc ou tatar. Le *bouriate* est parlé par deux cent mille individus environ, aux alentours du lac Baïkal, dans la Sibérie du sud. On trouve encore d'autres idiomes mongols aux environs de Kaboul.

Bien que ces idiomes offrent tout autant d'intérêt que ceux du groupe précédent, nous n'en dirons que quelques mots rapides. L'ensemble de leurs caractères est très analogue, en effet, à tout ce que nous avons vu dans le cours de ce chapitre ; c'est dans le vocabulaire ou dans le degré de développement grammatical que se trouvent les principales différences existant entre le tongouse et le mongol.

Le mongol possède un alphabet assez analogue à celui du mandchou, et dérivant comme lui (ainsi que l'alphabet kalmouk) de l'écriture syrienne. Il comprend sept voyelles, *a*, *e*, *i*, *o*, *u* (« ou » français), *eu* français, *ü* (notre « u »), et dix-sept consonnes, parmi lesquelles on compte les sons *ts* et *dz* ; la forme des caractères varie selon qu'ils se trouvent placés au milieu, au commencement ou à la fin des mots. Chaque consonne, ainsi que cela se passe dans l'alphabet hindou, est accompagnée d'une voyelle, sauf dans le cas où cette consonne est terminale.

L'harmonie vocalique, qui caractérise les idiomes ouralo-altaïques, se retrouve dans les langues du groupe mongol, mais avec quelques particularités. Parmi les faits phonétiques propres à cette famille, on peut relever, en bouriate, l'élision des voyelles finales et certaines modifications qu'éprouvent les consonnes lorsqu'elles sont en contact.

D'autre part, remarquons qu'en mongol les pronoms régimes ne sont pas incorporés, à la différence de ce qui se passe dans presque toutes les autres langues ouralo-altaïques. Tandis, par exemple, que le turc rend en un seul mot ces expressions « je le vois, je le mange », le mongol les rend par deux mots.

Le bouriate, l'obscur bouriate, joue dans la classe mongole un rôle très important ; selon L. Adam, le développement grammatical du bouriate est d'autant plus

instructif que l'on peut y reconnaître les formes intermédiaires par lesquelles ont passé les pronoms pour devenir des suffixes. Ce phénomène de la supériorié d'un idiome pour ainsi dire sauvage sur des langues littéraires et cultivées comme le sont le mongol et le mandchou, est loin d'être rare.

VI. *De l'harmonie vocalique et de la parenté des langues ouralo-altaïques.*

Le phénomène de l'harmonie des voyelles dans les langues altaïques est d'autant plus important qu'il constitue un des principaux arguments sur lesquels on s'appuie d'habitude pour affirmer la parenté du samoyède, du finnois, du turc, du tongouse, du mongol. Qu'est-ce donc que l'harmonie vocalique, quel est son caractère, quelle est son origine, quelle est sa valeur, quelles conclusions peut-on tirer de son existence simultanée dans les différents idiomes dont il s'agit ?

Le fait auquel se réduit l'harmonie vocalique est celui-ci : les différentes voyelles étant réparties en deux classes, toutes les voyelles d'un mot qui suivent celle de la syllabe principale doivent être de la même classe que la voyelle de cette syllabe. Dans certains idiomes ouralo-altaïques, il y a pourtant des voyelles dites « neutres » qui vont indifféremment avec l'une ou l'autre classe. Voici d'ailleurs ci-contre le tableau de la classification des voyelles dans un certain nombre de langues ouralo-altaïques (1).

Dans ce tableau, le signe *u* représente notre « ou » français, le signe *ö* vaut « eu », le signe *ü* vaut « u » de « surplus ». La classification est à peu près la même

(1) L. Adam. *De l'harmonie des voyelles dans les langues ouralo-altaïques*. Paris, 1874.

partout. Les trois voyelles primitives *a*, *u*, *i* sont, en principe, les deux premières fortes, la dernière neutre ; les voyelles intermédiaires sont faibles, en principe également.

	VOYELLES		
	fortes.	faibles.	neutres.
En suomi.....	u, o, a	ü, ö, ä,	e, i
— magyar ...	u, o, a	ü, ö	e, i
— mordvin ..	u, o, a	ä, i	»
— zyriène ...	ô, a	ä, i, e	»
— turc	u, o, a, e	ü, ö, e, i	»
— mongol ...	u, o, a	ü, ö, ä	i
— bouriate ..	u, o, a	ü, ö, ä	e, i
— mandchou..	ô, o, a	e	u, i

Mais dans la pratique il règne une assez grande variété : l'harmonie peut s'étendre au mot entier ou être restreinte aux suffixes ; elle peut s'appliquer à tous les mots ou n'affecter que les mots simples, ceux qui ne sont pas composés. En turc, par exemple, tout le mot doit être harmonique, de même qu'en mandchou, en mongol, en suomi, en magyar, tandis qu'en mordvin et en zyriène les seules voyelles sensibles sont les voyelles des désinences.

En magyar les mots composés conservent leurs voyelles originaires.

Nous avons à nous demander quelle est la cause de ce phénomène.

L'harmonie vocalique des langues ouralo-altaïques est-elle primitive, est-elle récente ?

L. Adam, qui s'est préoccupé spécialement de cette grave question, réfute aisément l'opinion des auteurs qui ne voient dans l'harmonie, comme le fait O. Bœhtlingk, que le résultat de circonstances physiologiques locales, ou qui n'en font, comme F. Pott, qu'un accident purement mécanique.

Schleicher, et après lui A. Riedl, ont trouvé la véri-

table solution du problème. Schleicher ne s'était occupé que du cas le plus général et le plus remarquable, celui de l'harmonisation des voyelles des suffixes, qui présentent chacune aux yeux de l'étranger une double forme (forte au faible, suivant la nature de la voyelle qu'on y admet). Il avait été frappé de la circonstance de cette infériorité, de cette subordination de la désinence par rapport au thème, et il en avait conclu qu'elle était le résultat nécessaire de l'agglutination et de la tendance à juxtaposer le plus possible dans le langage la signification et la relation si étroitement unies dans la pensée. A. Riedl a montré qu'il en était bien ainsi, car l'étude des plus anciens monuments de la langue magyare révèle à cet égard, du douzième siècle jusqu'à nos jours, de très grands progrès. Dans les textes anciens, les formes antiharmoniques abondent : on trouve *halâl-nek* « à la mort », qui devrait aujourd'hui se prononcer *halál-nak*.

L. Adam en conclut très justement qu'antérieurement au douzième siècle, le nombre des mots dérivés avec assimilation vocalique était encore plus restreint et qu'ils étaient remplacés par de véritables composés inharmoniques. Soient, dit-il, deux radicaux : *fa*, arbre, et *vel* (*veli*), compagnon : *favel*, arbre-compagnon, sera le composé inharmonique de ces deux éléments nominaux, également doués d'une signification propre. Mais, après que *vel* aura été successivement juxtaposé à un certain nombre de radicaux, il tendra à manifester, d'une façon sensible, qu'il perd en se composant sa signification originelle de « compagnon », et qu'il joue par rapport au radical, placé en tête, le rôle d'exposant de la relation « avec » (1). Nous avons donc affaire à un phénomène de décadence provenant de l'oubli du sens primitif des syllabes dérivatives.

(1) *Op. cit.*, p. 67.

L'élaboration en a d'abord été lente ; elle a été inégale dans les divers idiomes ouralo-altaïques, dont plusieurs, comme le tchérémisse des montagnes et le viotaque, ne connaissent que les rudiments de l'harmonie vocalique. Mais L. Adam suppose que, sous de puissantes influences étrangères, ces derniers idiomes ont perdu l'harmonie ; selon lui, il resterait chez eux des traces suffisantes d'harmonie pour permettre de conclure qu'à une certaine époque toutes les langues ouralo-altaïques étaient soumises à l'harmonisation des voyelles. Certes, nous regardons l'harmonie des voyelles comme un caractère de grande importance ; mais, après tout, ce n'est qu'un fait historique et relativement récent. Nous n'avons pas à rechercher ici les causes et les conditions de son développement, mais nous pensons qu'à lui seul il ne saurait prouver la descendance commune des cinq groupes linguistiques dont nous venons de parler.

En définitive, nous pouvons dire que si la parenté des cinq langues ouralo-altaïques n'est pas définitivement démontrée, elle est au moins fort vraisemblable. Le jour n'est peut-être pas éloigné où paraîtra une véritable grammaire comparée ouralo-altaïque. On a commencé naturellement par étudier les divers idiomes de cette famille ; nous avons cité le nom de Castrén, celui de W. Schott ; nous devons y joindre celui de A. Schiefner. On doit au premier des travaux d'une importance capitale sur les dialectes samoyèdes, sur le zyriène, l'ostiaque, le bouriate, le tongouse. Ahlqvist, Budenz, Donner, Hunfalvy, Weske, d'autre encore, ont puissamment aidé aux progrès de la comparaison des différents idiomes (1). Quelques pas de plus, et il sera possible de songer à un

(1) Budenz. *Ugrische sprachstudien*. Pest, 1869. — Donner. *Vergeichendes wörterbuch der finnisch-ugrischen sprachen*. Helsingfors, 1874. — Weske (*aus Livland*). *Untersuchungen zur vergleichenden grammatik des finnischen sprachstammes*. Leipzig, 1872.

rapprochement en règle des cinq groupes principaux.

En tous cas, que l'on arrive à les ramener à une origine commune, ou que l'on échoue dans cette tentative (soit en ce qui les concerne tous, soit en ce qui ne concerne que quelques-uns d'entre eux), il est certain qu'on ne pourrait les séparer les uns des autres dans l'étude générale des langues agglutinantes, tant ils se trouvent rapprochés par les phénomènes de l'harmonie vocalique.

§ 15. La langue basque.

Cette langue si remarquable, si intéressante, n'est guère parlée aujourd'hui, en Europe, que par 450 000 individus, sans grande originalité sociale, sans existence politique distincte. Les trois quarts d'entre eux, environ, appartiennent à la nationalité espagnole, et 140 000 approximativement à la nationalité française. Sur les rives de la Plata on compte à peu près 200 000 Basques émigrés.

Il est bien entendu que nous ne parlons ici que des individus parlant la *langue* basque et que la question spéciale de la *race* reste en dehors de cette statistique. Nous savons, en effet, grâce aux excellents travaux de P. Broca, qu'il y a Basques et Basques ; que les Basques espagnols, par exemple, sont de sang bien moins mélangé que les Basques français.

On a longtemps cherché à tracer les limites de la langue basque ; c'est seulement en ces derniers temps que l'on est arrivé à des résultats méritant une véritable confiance. La carte la plus complète est celle du prince Bonaparte, imprimée en 1869, et qui n'a été publiée que plus tard. Broca, de son côté, a dressé une carte moins considérable (*Sur l'origine et la répartition de la langue basque ; Revue d'anthropologie* ,t. IV, pl. 3). On trouve

entre ces deux cartes de très petites différences. La liste des localités frontières a été donnée par J. Vinson, dans l'Avant-propos de l'*Essai sur la langue basque*, de Fr. Ribary. Partant un peu au-dessous de Biarritz, la limite linguistique rejoint l'Adour au bas de Saint-Pierre d'Irube (à 2 kilomètres au sud-est de Bayonne), suit ce fleuve jusqu'au-delà d'Urcuit, le quitte alors pour englober Briscous et Bardos (à l'exclusion de la Bastide-Clairence), puis Saint-Palais et Esquiule (près d'Oloron), pour aboutir au pic d'Anie. En Espagne, la limite du pays basque s'étend jusqu'en dehors de la vallée de Roncal (vers l'Aragon) ; après Burgui, elle s'infléchit à gauche vers Pampelune, qu'elle contourne extérieurement, pour redescendre jusqu'au-delà de Puente la Reina, et revenir ensuite, presque en ligne droite, à Vitoria, d'où elle remonte vers la mer, qu'elle atteint un peu à l'ouest de Portugalete.

Le pays basque se compose donc de la province espagnole de la Biscaye presque tout entière, du Guipuzcoa, de la partie septentrionale de l'Alava et de près de la moitié de la Navarre ; il comprend, en outre, en France, une commune de l'arrondissement d'Oloron, celui de Mauléon et celui de Bayonne presque intégralement ; ce qui correspond aux anciennes divisions locales de la Soule, de la Basse Navarre et du Labourd.

Il n'y a pas de preuve historique, de preuve vraiment historique, que le basque ait occupé dans les temps anciens une aire géographique plus étendue. Nous reviendrons plus loin sur la question dite *ibérienne*. En France on ne peut démontrer avec certitude que le basque ait été en usage dans aucun des villages où le gascon se trouve aujourd'hui employé de façon exclusive. Par contre, il est incontestable qu'en Espagne il a perdu du terrain depuis plusieurs siècles : Pampelune, tout espagnole aujourd'hui, était basque jadis. Et de nos jours il

est facile de constater une altération très sensible de la langue basque dans les localités un peu importantes où se fait plus vivement sentir l'activité de la vie moderne et où le contact est plus fréquent avec les étrangers. Le langage de Saint-Sébastien, par exemple, et celui de Saint-Jean de Luz sont particulièrement incorrects, et l'on y rencontre un grand nombre de mots espagnols ou français.

La zone mixte basque et espagnole comprend, entre autres localités, Bilbao, Orduna, Aoiz, Roncal. Dans son mémoire sur la répartition de la langue basque, P. Broca a expliqué ingénieusement l'absence d'une zone analogue entre le basque et le gascon. « En Espagne, dit-il, le basque se trouve aux prises, sur sa lisière avec le castillan, dans des conditions d'infériorité qui rendent inévitable l'empiètement graduel de cette dernière langue. Mais en France, la langue qui entoure le basque n'est pas comme le castillan, une langue officielle, administrative, politique et littéraire ; ce n'est qu'un idiome populaire, un vieux patois qui n'a aucune force expansive, qui est, au contraire, en voie d'extinction. Il n'y a aucune raison pour que ce patois supplante le basque, ni pour que le basque empiète sur lui. Les deux idiomes restent donc stationnaires, égaux dans leur faiblesse et menacés l'un et l'autre par le français, qui les remplacera tôt ou tard. La langue que les Basques ont intérêt à apprendre, c'est le français. Tous ceux qui ont quelque instruction le connaissent déjà ; tous les habitants des villes de quelque importance le parlent ou le comprennent. Chaque ville, chaque bourg deviendra ainsi un foyer de diffusion ; il arrivera un moment où le basque ne sera plus parlé que dans les hameaux les plus isolés et dans les vallées les moins accessibles, et là même il finira par tomber en désuétude. Il périra donc sous l'influence d'une cause qui, sans doute, n'agira pas sur tous

les points avec la même rapidité, mais qui agira partout à la fois. On ne le verra pas reculer pas à pas, comme il fait en Espagne, où le castillan l'envahit de proche en proche, car il n'est pas plus menacé sur sa lisière que dans le reste de son territoire.

« Il n'est pas dit, toutefois, que le basque doive se maintenir jusqu'à la fin dans ses limites actuelles. Il est assez probable que le patois béarnais qui l'entoure disparaîtra avant lui, et qu'alors le français, venant presser directement sur la frontière basque, la refoulera peu à peu vers le sud, c'est-à-dire vers les Pyrénées, dont les hautes vallées seront probablement le dernier refuge de la plus ancienne langue de l'Europe. » (*Op. cit.*).

Le nom propre, le nom original du basque est *escuara*, *euscara*, *uscara*, suivant les dialectes, d'où, en français, « euscarien », synonyme de « basque ». Les Espagnols donnent à la langue basque le nom de *vascuence*, à ceux qui la parlent le nom de *Vascongados*. Nous ne saurions nous prononcer sur l'origine de ces différents mots. En ce qui concerne *escuara*, l'étymologie la plus probable est sans doute celle de de F. Mahn ; ce mot signifierait « manière de parler », « langage ». Les explications que l'on obtient des Basques eux-mêmes, à ce sujet, sont des plus fantaisistes, ce dont nous ne devons pas nous étonner. Lorsqu'ils comparent leur langue aux idiomes des peuples qui les entourent, les Basques se trouvent si complètement désorientés, qu'ils tombent aussitôt dans la plus naïve admiration pour leur parler maternel. L'un d'eux, le jésuite Larramendi, dont le livre porte ce titre présomptueux : « El impossible vencido » — l'impossible vaincu — fait à peu près du basque la source commune de toutes les autres langues ; un autre, l'Astarloa, affirme que chacune des lettres de l'escuara possède une valeur mystérieuse ; un troisième, l'abbé Darrigol, démontre, à l'aide de Beauzée, l'éternelle perfection de

ladite langue ; Chaho invente son ingénieuse théorie des « voyants » basques, dont la civilisation précoce a été étouffée par les Celto-Scythes barbares ; l'abbé d'Iharce de Bidassouet fait de l'escuara la langue dont se servit le Père éternel pour converser avec le premier des Juifs. Est-il quelque folie à laquelle n'ait donné prétexte ce précieux débris des idiomes de l'Europe antique ?

A la vérité, l'escuara offrait des difficultés d'étude insurmontables à ceux qui n'étaient habitués qu'à commenter les textes grecs et latins au moyen de procédés empiriques ; aussi les savants du moyen âge regardaient-ils volontiers la langue basque comme une énigme indéchiffrable, comme un problème insoluble. Un proverbe conservé dans le nord de l'Espagne prétend que le diable lui-même demeura chez les Basques sept longues années durant sans parvenir à entendre un seul mot de leur langue. Ainsi s'explique cette remarquable définition d'un dictionnaire espagnol : VASCUENCE. *Lo que está tan confuso y oscuro que no se puede entender.* « BASQUE. Ce qui est si confus et obscur qu'on ne le peut entendre. »

Par malheur, beaucoup d'érudits fort peu linguistes, beaucoup d'amateurs étrangers, ont voulu résoudre le problème, sans préparation spéciale, et leurs efforts infructueux n'ont fait qu'exalter l'infatuation qu'inspirait déjà aux Basques le spectacle de tant d'efforts stériles. L'on a pu dire, non sans une certaine apparence de vérité, que l'étude du basque menait droit à la folie. Mais les choses sont bien changées aujourd'hui qu'il existe une méthode linguistique. Le sphinx, mieux attaqué, a livré son secret, et bien qu'il reste encore nombre de points à éclaircir, il est présumable que le jour n'est pas éloigné où l'on pourra se féliciter de connaître à fond les lois nombreuses et complexes de la langue basque. Il y avait assurément de bonnes, d'excellentes choses dans les écrits d'Oihenart, de Chaho, et surtout de Lécluse ;

mais les travaux récents (L.-L. Bonaparte, van Eys, Vinson) ont fait faire à la question des progrès décisifs.

Le basque, pour le savant qui l'étudie sans que ce soit sa langue maternelle, est dans un état complet d'isolement. Aucun des idiomes qui l'environnent ne peut, s'il s'agit de la formation des mots, de la morphologie, lui être comparé, et la langue qui lui ressemble le plus par quelques traits généraux, à savoir le magyar, s'en trouve géographiquement fort éloignée. L'histoire du magyar est d'ailleurs connue en partie, tandis que l'on ne sait rien de celle du basque. Impossible de rencontrer dans aucun monument authentique plus ancien que le dixième siècle des traces évidentes de la langue basque; encore ne peut-on faire remonter à cette époque qu'une charte latine datée de 980, qui délimite le diocèse épiscopal de Bayonne et donne les noms plus ou moins altérés de quelques localités du pays basque. Il est avéré que les prétendus chants de guerre euscariens attribués à un âge plus ancien — à plusieurs centaines d'années, disait-on, avant le dixième siècle — ne sont rien moins qu'apocryphes.

Du dixième au seizième siècle (sauf une liste de vingt mots dans un manuscrit de la fin du douzième, conservé à Saint-Jacques de Compostelle), rien encore que des noms épars dans diverses chartes, coutumes, lettres patentes, bulles pontificales. C'est Lucius Marinæus Siculus qui parla le premier de l'escuara dans ses « Cosas memorables de España » (Alcala, 1530), et qui, le premier, en cita quelques mots. Quant au plus ancien texte basque imprimé (le plus ancien du moins à notre connaissance), c'est le discours de Panurge dans le chapitre neuvième du deuxième livre de Rabelais. C'est en 1542 que ce morceau fut publié. Le premier livre imprimé est daté de 1545, ce sont les poésies moitié religieuses, moitié érotiques de Bernard Dechepare, curé de Saint-

Michel-le-Vieux, en Basse-Navarre, fidèlement réimprimées il y a quelque temps par l'éditeur Cazals, de Bayonne. Mais l'ouvrage le plus important pour l'étude du basque, c'est la version du Nouveau Testament par Jean de Liçarrague (de Briscous), ministre protestant à la Bastide-Clairence, imprimée à la Rochelle, en 1571, aux frais du Parlement de Navarre, par ordre de Jeanne d'Albret.

On peut dire sans doute que les changements subis par la langue depuis cette époque sont assez sensibles; mais, certainement, ils ne sont point considérables.

On constate même aujourd'hui des différences plus importantes entre les divers dialectes. Les variétés de la langue basque sont, pour ainsi dire, innombrables, et chaque village a quelque particularité qui lui est propre, Ce fait n'a sans doute rien d'anormal.

Mais, à côté du langage spontanément parlé, du langage local, les langues ont un dialecte général, en quelque sorte conventionnel, fruit de l'éducation, et qui souvent est très voisin de la langue écrite. Or, en basque, rien de tel, et chaque écrivain se fabrique un langage à sa fantaisie. Le prince L.-L. Bonaparte a compté jusqu'à huit dialectes qui ne formeraient pas moins de vingt-cinq variétés. C'est, en Espagne, le biscayen, le guipuscoan, le haut-navarrais méridional et le haut-navarrais septentrional. En France, le labourdin, parlé dans l'ancien Labourd (partie sud-occidentale de l'arrondissement de Bayonne) ; le souletin, dans les deux cantons sud-est de l'arrondissement de Mauléon (ancienne Soule) ; le bas-navarrais occidental et le bas-navarrais oriental, en usage dans la Navarre française, c'est-à-dire dans le reste de ces deux arrondissements.

Au surplus, ces huit dialectes se réduisent sans peine à trois grands groupes. Le premier de ceux-ci, formé du BISCAYEN, est particulièrement remarquable par l'origi-

nalité de son verbe ; le second se compose du SOULETIN et du BAS-NAVARRAIS : aspiration fréquente et variation des *u* en *i* : le troisième, aux formes généralement plus pleines et moins altérées que les formes du second groupe, comprend les quatre dialectes : GUIPUZCOAN, LABOURDIN, HAUT-NAVARRAIS du nord et du sud. Nous n'entreprendrons pas d'indiquer ici les différences plus ou moins notables qui distinguent ces dialectes les uns des autres; disons simplement que les quatre dialectes de France possèdent l'aspiration, dont sont totalement dépourvus les dialectes d'Espagne. Ajoutons, d'ailleurs, qu'en ce qui concerne l'intérêt spécial que peuvent offrir ces différents dialectes, le souletin, le labourdin, le guipuzcoan et le biscayen, sont à peu près les seuls étudiés, parce que ce sont les seuls qui aient eu une certaine littérature. Les dialectes du centre, guipuzcoan et labourdin, paraissent être les moins altérés, tandis que les deux autres ont subi chacun de plus profondes modifications, en tout cas J. Vinson donne le pas au labourdin même sur le guipuzcoan.

On comprend aisément que c'est par l'étude simultanée et comparative des huit dialectes qu'il est possible de déterminer le caractère général de la langue basque, en reconstituant, autant que faire se peut, ses formes communes. La phonétique peut seule conduire à ce résultat. Jetons donc un rapide coup d'œil sur la phonétique du basque.

Le basque compte cinq voyelles simples : *a*, *e*, *i*, *o*, *u* ; *six* diphthongues ; *ai*, *ei*, *oi*, *ui*, *au*, *eu* ; les deux demi-voyelles *y* et *w*, et vingt-deux consonnes que l'on peut classer ainsi : *k*, *g*, *kh*, — *tch*, *ts*, — *t*, *d*, *th*, — *p*, *b*, *ph*, — trois *n* (*n* du grec ἄγγελος, *n* mouillé de « régner », *n* dental) *m*, — les sifflantes *h*, *ch*, *z*, *s*, — *r* dur (presque deux « rr »), *r* doux (très voisin de « l »), enfin *l*. Mais, s'il fallait comprendre dans cette liste les sons particu-

liers aux différents dialectes, elle serait plus que doublée ; il faudrait y faire entrer notre « u » français, propre au souletin, notre « j » la jota espagnole, et des « g », « t », « d », « l », mouillés. Ajoutons que le double *w* a la valeur du « w » anglais, mais il n'est qu'euphonique ; il n'apparaît qu'entre *u* et *a* ou *e* qui le suivent immédiatement.

Les lois phonétiques du basque sont assez nombreuses, et il n'est pas sans intérêt de signaler ici quelques-unes des plus importantes d'entre elles.

Deux voyelles se trouvent-elles en présence, la première est élidée si elle se trouve à la fin d'un mot ; que la rencontre, au contraire, ait lieu dans le corps même d'un mot, l'hiatus est de règle, mais l'une des voyelles doit changer : *e* devient *i*, *o* devient *u* ; toutefois *a* persiste. Les variations que subissent les consonnes, lorsqu'elles se rencontrent, sont bien plus remarquables : une consonne dure, placée à la fin d'un mot, rencontre-t-elle au commencement du mot suivant une consonne douce, la dure disparaît et la douce devient dure ; c'est ainsi, par exemple, que *hunat goiti*, « ici en haut », se prononce *hunakoiti*. Autres lois : les explosives dures (par exemple *k*, *t*) tombent devant les nasales ; après une sifflante, les explosives doivent être dures ; après une nasale, elles doivent être douces. Une consonne ne peut être doublée, c'est-à-dire que deux g, par exemple, deux t, et ainsi de suite, ne peuvent se suivrent immédiatement ; les explosives dures initiales deviennent douces spontanément ; entre deux voyelles simples, *g*, *d*, *b*, *n* et *r* doux sont absolument supprimés ; aux mots étrangers l'on préfixe une voyelle : c'est ainsi que « raison » devient *arrazoin*.

L'orthographe la plus généralement admise aujourd'hui dans le pays basque est assez nouvelle et n'est d'ailleurs qu'une réforme des vieux usages du pays. Le

basque n'ayant point conservé — s'il en a jamais eu, ce qui n'est prouvé en aucune façon — de caractères graphiques particuliers, il avait fallu, pour l'écrire, adopter l'alphabet latin tel que l'employaient les Gallo-Romains ou Hispano-Romains de la région des Pyrénées. Deux orthographes, sensiblement différentes, se trouvèrent en présence, l'une française, l'autre espagnole. Chacune avait ce défaut capital de représenter souvent le même son, un seul et même son, par des lettres différentes ; on écrivait, par exemple, *z*, *c*, *ç* pour *s*, et *c*, *qu*, *k* pour *k*. L'orthographe réformée s'inspira du système espagnol plutôt que du système français ; toutefois *z* s'y prononce *s* (« s » de « sur, sœur, son »).

Arrivons maintenant à la formation des mots.

La soi-disant déclinaison du basque consiste en simples particules placées à la suite du mot. L'escuara ne dit pas, par exemple, « à la femme » ; il dit : « femme la à » ; au lieu de prépositions, il emploie des postpositions, c'est-à-dire des suffixes plus ou moins agglutinés au nom ou à l'article. Les principaux suffixes dont il s'agit sont en « de », indiquant le génitif ; *i* « à », indiquant le datif ; *ko* « de, pour » ; *tik* « de », indiquant l'ablatif ; *n* « dans », *z* « par », *kin* ou *gaz* « avec », *ra* « vers », *ik*, ayant le sens positif de l'anglais « some » (c'est-à-dire le sens de notre « de » français dans les phrases comme celle-ci « donnez-moi de l'eau »), *no* « jusque », *gabe* « sans », *gatik* « à cause de », *tzat* « pour », etc. On donne le nom de déclinaison définie à celle qui comporte l'article. Les grammairiens distinguent encore dans leurs tableaux la déclinaison des êtres raisonnables d'avec celle des êtres dépourvus de raison : la première serait caractérisée par l'intercalation, entre l'article et le suffixe, d'une syllabe particulière, *baith* — syllabe inexpliquée, mais que les étymologistes ont naturellement assimilée sans hésitation à l'hébreu *beth*

« maison », vue que cette intercalation n'a lieu qu'auprès des suffixes locaux, « dans, chez, vers », etc. — La déclinaison indéfinie est unique, c'est-à-dire qu'elle n'a en quelque sorte ni pluriel ni singulier ; le basque, en effet, ne peut joindre à ses noms le signe du pluriel, quand ils ne sont pas déterminés : il ne saurait dire « femmes », force lui est de dire « les femmes ». Il en résulte que le signe du pluriel (qui est *k*) se place seulement après l'article *a* (ancien pronom démonstratif conservé en biscayen). La déclinaison définie a donc, grâce à cet article, un singulier et un pluriel.

Il est bon de noter que de grandes irrégularités signalent l'adjonction des suffixes au nom : quelquefois, par exemple, l'article et le signe du pluriel disparaissent. Mais ce sont là des détails particuliers que nous devons négliger dans une esquisse aussi sommaire que celle-ci.

L'on conçoit aisément, d'après ce qui précède, à quel point il serait inexact de se servir pour le basque des mots de cas, de nominatif, de génitif, et ainsi de suite. On emploie parfois ces expressions, mais il ne faudrait point s'y laisser tromper ; ce ne peut être qu'une façon de parler. En basque, il n'y a pas et ne saurait y avoir de suffixe nominatif, accusatif ou autre — par exemple les *s*, *m* du latin « dominu-s, dominu-m » : — l'on se sert du thème pur et simple, du thème tel quel ; seulement ce thème, lorsqu'il est sujet d'un verbe actif, se voit postposer un suffixe *k* dont l'origine est inexpliquée :*gizonak eman du* « l'homme l'a donné », *gizonek yo dute* « les hommes l'ont frappé » ; *gizon* veut dire « homme », *a* est l'article, *k* le signe du sujet dont nous venons de parler.

Une particularité de la langue basque qui ne laisse pas que d'étonner, c'est le grand nombre de mots, souvent réduits à une syllabe, qui viennent s'annexer à d'autres mots pour marquer l'augmentation, la diminution, l'abondance, la mauvaise qualité, l'excès, le défaut, l'atta-

chement, la répugnance, et ainsi de suite. Mais beaucoup de nos langues modernes n'ont-elles pas (à un moindre degré, il est vrai) la faculté de former des diminutifs et des augmentatifs ?

L'adjectif, toujours invariable, se place constamment après le nom ; le basque, pour rendre la phrase que voici : « la belle maison du petit homme », devra dire : « homme petit le de maison belle la ». Remarquons que l'adjectif s'intercale ici, comme chez nous, du reste, entre l'article et le nom auquel il se rapporte, mais que le génitif (« de l'homme ») précède le nom (« la maison ») dont il dépend.

Les pronoms personnels sont *ni* « moi », *gu* « nous », *hi* « toi », *zu* « vous ». Le basque contemporain emploie comme nous la seconde personne du pluriel pour exprimer poliment le singulier ; aussi s'est-il fabriqué un « vous » pluriel : *zuek*. Point de pronoms relatifs ; pour imiter les Français ou les Espagnols, les Basques modernes ont employé souvent avec le sens relatif les pronoms interrogatifs ; mais cela est absolument contraire au génie même de leur langue.

En ce qui concerne les noms de nombre, nous pouvons remarquer que le basque ne possède pas de mot original pour exprimer « mille » et que tout, chez lui, indique une numération vigésimale : « trente-neuf » est pour lui vingt et dix-neuf, « soixante » est trois-vingts.

Le verbe basque est simple ou périphrastique. La conjugaison simple est caractérisée par ce fait que les éléments dérivatifs marquant le temps, le mode, la personne sont unis au radical ; la conjugaison périphrastique a lieu au moyen de deux verbes simples auxiliaires, « être » et « avoir »—*dut* et *naiz*—joints à un nom d'action décliné.

La question du verbe basque est d'une grande impor-

tance ; c'est elle qui déroute le plus les esprits accoutumés à nos grammaires classiques grecques et latines, et l'on ne peut dire encore que ce soit une question résolue. Pour Mahn, van Eys, Vinson, Ribary, la conjugaison simple est seule primitive et l'autre n'est qu'une composition dont l'origine ne remonte qu'à la période historique de la langue basque. Sans entrer dans les raisons spéciales qui militent invinciblement, selon nous, en faveur de cette opinion, nous ferons simplement remarquer que l'opinion opposée (contestant l'existence d'un radical à sens verbal dans les formes des auxiliaires) a un caractère métaphysique qui la rend inacceptable de prime abord.

La conjugaison basque périphrastique a cet avantage de permettre pour chaque verbe une double expression répondant au sens transitif et au sens intransitif : la voix intransive, c'est un nom d'action accompagné de *naiz* « être » ; la voix transitive, c'est un nom d'action accompagné de *dut* « avoir ». Nous verrons en temps et lieu que le verbe sémitique exprime le régime direct par un élément pronominal suffixé au verbe ; nous avons vu un peu plus haut que le magyar, le vogoul, le mordvin agissent de même en semblable occasion (sans toutefois placer l'élément en question au même lieu que le placent les langues sémitiques); or, le basque agit de même. Toutefois, il a cette infériorité sur les langues que nous venons de citer, qu'il ne saurait séparer du verbe actif son régime direct : il ne pourrait dire, par exemple : « j'aime une femme »; il ne peut dire que ceci : « je l'aime une femme ». Mais, dans son verbe, le basque exprime le régime indirect et dit en un seul mot : « je le donne à lui »; ici encore il ne peut omettre le régime direct, il ne peut dire : « je donne à lui ».

Chacune de ces formes complexes est susceptible de quatre modifications, suivant qu'on parle familièrement

à un homme ou à une femme, qu'on s'adresse à une personne que l'on veut honorer, ou qu'enfin l'on ne veuille pas tenir compte de ces circonstances. Les grammairiens désignent ces modifications sous le nom de traitement masculin, féminin, respectueux, indéfini.

Certains caractères de la langue basque se retrouvent, ainsi qu'on l'a souvent répété, dans les langues américaines. Le verbe basque a sans doute quelques analogies avec la conjugaison des langues de l'Amérique ; mais de là à conclure, comme le font sans hésiter quelques auteurs, à une parenté intime entre l'algonquin, l'iroquois, par exemple et l'escuara, il y a loin, fort loin. Avant d'affirmer que le basque est, comme ces idiomes, polysynthétique ou incorporant, il conviendrait de déterminer ce que c'est au juste que le polysynthétisme ou l'incorporation. Dans le paragraphe relatif aux langues américaines, nous essayerons de définir ces deux expressions.

Nous nous contenterons, pour l'instant, de signaler ici une particularité des idiomes du nouveau monde que l'on retrouve en basque, à savoir : la composition par syncope, qui, d'ailleurs, n'est point tout à fait inconnue aux langues européennes modernes : de *ortz* « nuage » et *azantz* « bruit », le basque fait *ortzanz* « tonnerre, bruit du nuage ». Mais les composés de cette espèce ne sont pas très nombreux. Pour l'ordinaire, on les rencontre dans les noms de lieux, ces restes précieux d'une époque antérieure et si souvent réfractaires à l'analyse.

Il se peut que les noms de lieux nous apprennent un jour bien des mots tombés en désuétude et finalement oubliés.

Dans son état actuel, et bien qu'il soit imparfaitement connu, on peut dire que le vocabulaire escuara est assez pauvre. Exclusion faite des nombreux mots gascons,

français, espagnols et latins qu'il renferme, et encore d'autres mots qu'il est possible de rattacher à quelque autre source, il est probable que les termes réellement basques n'expriment, en général, aucune idée abstraite. C'est ainsi que l'on ne connaît pas de mot basque simple ayant le sens général que nous attachons en français au mot « arbre », au mot « animal ». C'est ainsi encore qu'il existe en basque six expressions différentes pour exprimer l'état de chaleur de la chienne, de la jument, de la vache, de la truie, de la brebis et de la chèvre, mais il n'existe point, paraît-il, de terme général applicable à cet état d'une façon commune. Pour les Basques, « dieu » est « le seigneur d'en haut », et s'ils ont un terme pour exprimer notre mot de « volonté », ce terme signifie également « pensée, désir, fantaisie ».

Pour reconstituer, autant que faire se pourra, le vocabulaire commun escuarien, il s'agira de rechercher tous les mots usités dans les divers dialectes, et il conviendra, naturellement, de ne les admettre comme originaux qu'après certitude acquise qu'ils n'appartiennent point en propre à quelque idiome étranger. L'histoire nous apprend que la région où se parle l'escuara a été traversée par des peuples de langue celtique, par des Germains, par des Arabes, surtout par des peuples de langue romane. L'influence du latin a dû être d'autant plus puissante qu'elle a été continue vingt siècles durant et qu'elle s'est exercée plus activement ; pour bien connaître le basque, il importe donc de savoir à fond le latin et l'histoire de ses deux dérivés, le français et l'espagnol.

En ce qui concerne l'origine même des Basques, si l'idiome escuara a été la langue des ancien Ibères, ou du moins l'un des dialectes de leur langue, le fait n'est pas encore scientifiquement démontré. D'après de très anciennes traditions, les Ibères formaient le peuple qui,

avant l'arrivée des nations de langue indo-européenne, habitait l'ensemble de la péninsule ibérique, soit l'Espagne et le Portugal : ils occupèrent également, paraît-il, toute la partie de la Gaule qui plus tard reçut le nom de Narbonnaise. Leurs premiers rapports connus avec des individus de race étrangère remontent au temps des expéditions phéniciennes dont l'histoire nous a transmis le souvenir. Advint l'invasion celtique, qui donna naissance aux Celtibères. Ceux-ci résistèrent vaillamment aux légions romaines et supportèrent le choc des musulmans après avoir subi la domination des Visigoths. Conservé dans la région où vivaient les Ibères; l'escuara n'étant ni sémitique ni indo-européen, on fut amené à le tenir pour le représentant direct, au moins pour l'un des anciens représentants de la vieille langue ibérienne.

L'on s'appuie d'habitude, pour soutenir cette opinion, sur trois sortes de preuves, preuves tirée des mœurs, preuve tirée du type, preuve tirée de la langue.

L'argument tiré des mœurs se résume dans une disposition légale des coutumes françaises de la région pyrénéenne, même en dehors du pays basque, qui établissait dans les successions un droit absolu de primogéniture sans distinction de sexe ; or Strabon dit que chez les Cantabres, qui paraissent être une tribu ibérienne, les filles héritaient. J. Balasque a démontré que le droit de primogéniture provenait du principe celtique ou gallique de la conservation intégrale du patrimoine.

Le type est aujourd'hui connu. Nous possédons les caractères du véritable crâne basque, celui d'Espagne. Mais ce type aurait beau s'être étendu sur l'Espagne entière — nous le rencontrons sans doute aussi en Corse et dans le nord de l'Afrique — cela ne prouverait en aucune façon que cette seule et même race n'ait point parlé plusieurs langues différentes, comme cela se voit pour d'autres peuples.

Les preuves linguistiques se résument en essais d'explications de mots ibères par le basque. Les monuments de la langue ibérienne parvenus jusqu'à nous sont de deux sortes : d'une part des médailles et des inscriptions, d'autre part des noms propres et surtout des noms topographiques transcrits par des auteurs grecs et latins. Les médailles et les inscriptions offrent les éléments d'un alphabet dérivé du phénicien, mais il ne faudrait pas prendre le change sur leur prétendu déchiffrement : il n'est rien moins que certain. Nous ne pouvons admettre, dans les différentes lectures proposées jusqu'à ce jour, que des traductions aventureuses et forcées. La forme des noms recueillis par Strabon, Pline et autres anciens auteurs présente, au contraire, une base appréciable, mais il est tout naturel que les étymologistes en aient abusé, et largement, selon leur coutume. Les explications proposées par Humboldt, et, après lui, par nombre d'étymologistes sans principes, sans méthodes, sont très douteuses. Van Eys et Vinson, dont l'opinion est d'un poids considérable en la matière, sont d'accord à ce sujet (1).

Il est possible que les présomptions de Humboldt soient justes ; il est possible, peut-être même est-il vraisemblable et probable que les anciens habitants de l'Ibérie aient parlé une langue alliée au basque, sinon même une forme plus ancienne du basque ; mais que cela soit prouvé, nous nous refusons encore à l'admettre.

En résumé, cette prétendu identité est possible, mais les faits accumulés pour la faire accepter ne lui ont donné d'autre caractère que celui d'une hypothèse simplement plausible et qui attend encore sa justification.

(1) Van Eys, *la langue ibérienne et la langue basque* (*Revue de linguistique*, t. VII, p. 1. Paris, 1874) ; Vinson, *la Question ibérienne* (Mémoires du congrès scientifique de France, t. II, p. 357. Paris, 1874).

§ 16. Les langues américaines.

On a beaucoup écrit, on écrit beaucoup sur les langues américaines. Il en est peu qui aient prêté à autant de théories excentriques et fantaisistes ; elles partagent largement, sous ce rapport, le sort qui était réservé au basque et à l'étrusque.

L'idée commune, l'idée capitale qui obsède presque tous les américanistes est de rattacher les idiomes du nouveau monde à tel ou tel groupe des langues agglutinantes du monde ancien, le plus souvent aux langues ouralo-altaïques, parfois au basque, parfois au japonais, parfois à toutes les autres langues agglutinantes. Cela n'est pas sérieux. Avant de rapprocher l'iroquois du magyar, le totonak du japonais, le nahuatl du basque, il faudrait rapprocher scientifiquement les unes des autres toutes les langues américaines et esquisser au moins leur grammaire comparée. C'est ce que l'on ne se soucie point d'entreprendre.

On ne trouve nulle part comme en Amérique un nombre considérable d'idiomes se ressemblant autant les uns aux autres, mais constitués cependant au moyen d'éléments parfaitement divers. C'est pour ce motif que l'étude des idiomes du nouveau monde offre tant de difficultés et qu'il est assez embarrassant de se faire sur leur ensemble une idée bien précise.

A la vérité, il existe un grand nombre de grammaires, de livres de piété, de catéchismes, de versions de la Bible, propres à faciliter l'étude de certaines langues américaines ; mais la plupart de ces ouvrages ont été rédigés dans un esprit si peu scientifique et d'une façon si défectueuse, que toute l'aide qu'on en peut attendre

se réduit, en réalité, à fort peu de chose. Parmi les plus instructifs de ces écrits, nous signalerons celui de John Pickerings, *Remarks on the Indian Languages of North-America*, déjà ancien, le *Mémoire sur le système grammatical des langues de quelques nations de l'Amérique du Nord* de Duponceau, diverses notices de Mahn, Charencey, Fr. Müller (*Grundriss der Sprachwissenschaft*, t. II, 1re partie, p. 181), les récents écrits de L. Adam, de Brinton (1). Les *Etudes sur quelques langues sauvages de l'Amérique* par N. O***, ancien missionnaire, contiennent une esquisse intéressante et vraisemblablement très fidèle de l'algonquin et de l'iroquois, mais l'auteur s'y montre trop peu au courant des procédés méthodiques les plus élémentaires.

Dans la seconde édition de son Ethnographie, Fr. Müller répartit en 28 groupes les langues américaines :

1. Le groupe *kénaï*, au nord-ouest de l'Amérique septentrionale.

2. Le groupe *athapasque*, à l'est du précédent, s'étendant des bords du Youkon et du Mackenzie jusqu'à l'embouchure du Churchill dans la baie d'Hudson. On rencontre beaucoup plus au sud, et séparés du gros de ce groupe, d'autres idiomes qui s'y rattachent également : la langue des Qualihoquas au nord du fleuve Colombie, des Umpquas au sud ; celle des Apaches, plus au sud encore, dans l'Etat de Névada et la Californie supérieure.

3. Le groupe *algonquin*, au sud de la baie d'Hudson, et s'étendant à l'est jusque vers l'Atlantique. Il comprend le mikmak dans l'est de la Nouvelle-Bretagne et à Terre-Neuve ; la langue des Lenni-Lennapés ou Délawares (narraganset, mohican, etc.); le kri, l'ojibway, l'ottawa et d'autres idiomes encore.

(1) *The philosophic Grammar of american Languages* (1885). *On Polysynthesis and Incorporation* (1885).

4. Le groupe *iroquois* (onondago, sénéka, onéida, kayuga, tuskarora).

5. Le *dakota*, au centre de l'Amérique septentrionale, langue des Sioux et d'autres tribus.

6. Le *pani* ou pawnie.

7. Le groupe *appalache*, comprenant, entre autres idiomes, le chéroki, le kataba, le chakta, le krik, le natchez.

8. En remontant vers le nord-ouest, le *koloche* dans l'extrême ouest de la Nouvelle-Bretagne.

9. Le groupe *orégonais* plus au sud.

10. Le groupe *californien* (périkou, monki, kotchimi).

11. Le groupe *yuma* sur le bas Colorado.

12. Les idiomes indépendants des Pueblos de la Sonora et du Texas (zuni, tégua et autres).

13. Les idiomes indépendants du Mexique : le totonak, l'otomi, le taraska, le mixtek, le zapotek, le mazahua, le mame et autres.

14. Le groupe *aztek* et des idiomes sonoriens, comprenant d'une part le nahuatl ou aztek, et de l'autre un certain nombre des langues de la Sonora : kahita, kora, tarahumara, tépéguana ; opata, tubar ; pima, papago ; kizh, nétéla, kahuillo ; chochoni, komanche, moki, utah, pahutah, etc.

15. Le groupe *maya* dans le Yucatan, comprenant le maya au nord, le quiché, le huastek au nord-est de Mexico.

16. Les idiomes indépendants de l'Amérique centrale et des Antilles : le kuéva vers l'isthme de Panama, le cibuney dans les Antilles.

17. Le *caraïbe* et l'*arévak* ; le premier dans le Vénézuéla et la Guyane, le second dans la Guyane anglaise et la Guyane hollandaise (Brinton, *Transact. of the americ. philos.* Soc., XIV, p. 427).

18. Le *tupi*, le *guarani* et l'*omagua*. Les deux premiers

forment à eux deux un groupe plus particulier ; ils comprennent des idiomes parlés dans la région du Parana, du Paraguay, de l'Uruguay.

19. Certains idiomes de l'intérieur du Brésil, par exemple la langue des Botocudos.

20. Diverses langues de la Colombie.

21. Les langues indépendantes de la région des Andes.

22. L'*araucan*.

23. Le *guaykuru*, parlé entre le Paraguay et le Pilcomayo ; l'*abipon* dans la région du Saldo, au centre de la Plata.

24. Le *puelche* dans les Pampas, à l'ouest de Buenos-Ayres.

25. Le *téhuelche*, langue des Patagons, au sud du précédent.

26. Les différents idiomes de la Terre de Feu.

27. Le *chibcha*, de l'autre côté des Andes, dans la Colombie ou Nouvelle-Grenade, jusqu'aux environs de Santa-Fé de Bogota.

28. Le groupe *quichua* est parlé plus au sud, depuis la limite des Etats de Colombie et de l'Equateur jusque vers le tiers septentrional du Chili. Parents des Quichuas, les Aymaras sont sur la limite du Pérou et de la Bolivie.

Tous ces idiomes passant pour se ressembler et pour avoir un aspect général identique, nous avons à déterminer maintenant quel peut être ce caractère commun.

Et tout d'abord, leur mode d'être, leur mode de fonctionner sont-ils tellement étranges, tellement particuliers, qu'il faille renoncer à les classer dans une de ces trois grandes catégories, langues isolantes, langues agglutinantes, langues à flexion, sous lesquelles viennent se grouper les idiomes de l'ancien monde ?

C'est ce qu'un certain nombre d'auteurs ont pensé. Les langues américaines auraient, aux yeux de ces auteurs, une propriété spéciale qui suffirait à constituer une

classe bien à part, un quatrième système, qu'il faudrait appeler le système *incorporant* ou *polysynthétique*.

En tâchant d'éviter autant que faire se pourra les détails trop arides, recherchons et examinons quels sont les phénomènes sur lesquels a pu s'appuyer cette doctrine d'une classe incorporante, quelle est la nature de ces phénomènes et quelle est leur importance.

Nous terminerons par un coup d'œil rapide sur les idiomes algonquins et iroquois, parlés dans une grande partie de l'Amérique septentrionale et qui sont assurément les mieux connus de toutes les langues américaines.

A plusieurs reprises déjà nous avons dit ce qu'il fallait entendre par les termes de langue isolante, de langue agglutinante. Le monosyllabisme est caractérisé par l'emploi constant de racines indépendantes et invariables ; l'agglutination comporte l'état de dépendance de certaines racines, qui, vis-à-vis des autres racines dont le sens premier et entier a persisté, n'expriment plus que les idées du rapport, de relation. Nous verrons plus tard que la flexion ne doit être retrouvée que là où pour exprimer les divers rapports de temps et de lieu, la racine elle-même peut être organiquement altérée. Impossible de commettre une erreur sur la place à assigner à un idiome donné, si l'on a pu constater chez lui l'un de ces caractères, celui du monosyllabisme, celui de l'agglutination, celui de la flexion. Ainsi les langues sémitiques sont flexionnelles par excellence, bien qu'on y trouve l'agglutination — par exemple dans les préfixes et suffixes pronominaux du verbe et même dans le développement des voix dérivées ; — les idiomes indo-européens emploient bien souvent les procédés de l'agglutination ; mais, la flexion jouant un rôle essentiel dans la langue indo-européenne commune que l'on a pu restituer en partie, il faut les ranger forcément dans la classe des langues à flexion. H. Chavée avait donc un certain droit

à tenir pour défectueux le nom qui a été donné à la classe intermédiaire. En effet, quel que soit le degré de fusion, d'usure des éléments constitutifs, du moment qu'il y a autant de racines distinctes que d'idées principales et secondaires, subordonnées, attributives, il y a agglutination. A ce point de vue le sanskrit ne diffère en aucune façon du magyar. Nous traiterons dans notre sixième et dernier chapitre des empiètements d'une classe sur l'autre et de la certitude parfaite de la succession des trois états, monosyllabisme, agglutination, flexion.

Le nombre des idiomes agglutinants est considérable, mais l'agglutination revêt chez ces idiomes toutes les formes possibles. S'agit-il d'établir une division morphologique secondaire, une sous-classe, il ne faut donc pas se régler seulement sur le degré d'intensité de l'agglutination ou sur son abondance, il convient aussi de prendre en sérieuse considération l'ordre habituel des éléments mis en présence les uns des autres, c'est-à-dire la tendance plus ou moins marquée des idiomes en question à préfixer, suffixer à la racine, ou encore infixer les éléments formatifs des mots.

Telle était bien la pensée de Schleicher, lorsqu'il se refusait à reconnaître une quatrième catégorie de langues, composée des idiomes américains.

Qu'est-ce donc, en définitive, que ce « polysynthétisme », que cette « incorporation » que l'on prétendrait nous faire accepter comme caractérisant ce quatrième type de forme du langage ? Voici ce que dit à ce sujet Fréd. Müller dans son Ethnographie générale : « Les langues américaines reposent, dans leur ensemble, sur le principe du polysynthétisme ou de l'incorporation. En effet, tandis que dans nos langues les conceptions isolées que la phrase relie entre elles se présentent sous la forme de mots détachés, elles se trouvent réunies, au contraire, dans les langues américaines, en une indivisible unité.

Par conséquent, mot et phrase s'y confondent tout à fait. »

D'après les américanistes les langues qui nous occupent en ce moment auraient comme caractères distinctifs les particularités suivantes :

Premièrement, elles réuniraient au verbe des pronoms ou même des noms régimes.

Secondement, elles posséderaient une conjugaison nominale possessive.

Troisièmement, elles feraient varier le verbe lorsqu'il s'agit d'exprimer que l'objet de l'action a changé ou lorsqu'il faut nuancer l'action.

Quatrièmement enfin, elles connaîtraient le procédé de composition indéfinie par syncope et par ellipse.

Le premier et le second de ces prétendus caractères ne résistent même pas à un examen superficiel. Prenons d'abord l'union d'un nom et d'un pronom. Evidemment ce procédé est familier aux langues sémitiques et à bien des idiomes agglutinants de l'ancien monde. Lorsqu'avec le mot *mokkumân* « couteau » le kri forme les expressions *nimokkumân* « mon couteau », *kimokkumân* « ton couteau », *omokkumân* « son couteau », il court au procédé qu'emploie le magyar en disant *munkám* « mon ouvrage », *munkád* « ton ouvrage », *munkája* « son ouvrage » ; qu'emploie l'arabe en disant *kitabi* « mon livre », *kitabo* « son livre ». Il est vrai qu'en kri l'élément pronominal est placé avant le nom et qu'en magyar comme en arabe il est placé après le nom, mais cela n'a point d'importance en ce qui nous occupe en ce moment.

En ce qui concerne les variations du verbe destinées à nuancer l'action, Duponceau cite d'après Molina le chilien *elun* « donner », *eluguen* « donner plus », *éluduamen* « avoir le désir de donner », *eluzquen* « paraître donner », *eluvalen* « pouvoir donner », etc. ; mais cet

exemple ne rappelle-t-il pas très exactement les nombreux exemples analogues que l'on peut tirer du turc ? D'ailleurs, dans beaucoup de langues agglomérantes, l'on trouve des traces de pareilles dérivations, fort analogues, en somme, aux voix du verbe sémitique. Nous avons déjà cité des exemples tirés des langues dravidiennes et du basque.

L'on peut considérer comme un peu plus sérieuse la troisième des prétendues caractéristiques, à savoir cette circonstance que le verbe varie à mesure que varie l'objet de l'action. En chéroki, par exemple, on dit *kutuwo* « je me lave » *kukûsquô* « je me lave la figure », *tsêkûsquô* « je lave la figure d'un autre », *takungkalâ* « je lave mes vêtements », *takutêyâ* « je lave des plats », etc., en termanacan l'on dit *jucurù* « manger du pain », *jemeri* « manger du fruit, du miel », *janeri* « manger des aliments cuits », etc. ; en lénapé on emploie des verbes différents pour dire « manger de la soupe » et dire « manger de la bouillie ». Mais, en définitive, n'est-ce pas là une composition syncopée ? En ce cas nous retombons dans un caractère que nous aurons à examiner tout à l'heure. S'il n'en est pas ainsi, nous ne pouvons voir dans ce phénomène que cette horreur de l'abstraction, cette absence d'idées générales déjà constatée dans de nombreux idiomes de la catégorie agglutinante.

La réunion des pronoms régimes au verbe s'opère par des procédés analogues à celui de la conjugaison nominale ; aussi est-elle généralement pratiquée chez les peuples qui joignent aux noms les affixes possessifs. Le basque présente à ce sujet une exception remarquable : il ne connaît pas les affixes nominaux. Par contre, sa conjugaison « objective » est la plus riche de toutes celles des langues européennes et asiatiques ; il incorpore, en effet, dans le verbe, non seulement les pronoms régimes directs — « moi, toi, lui » — mais aussi les régimes in-

directs, tandis que le mordvin — idiome du groupe ouralo-altaïque finnois — ne sait exprimer que le régime direct des trois personnes ; le vogoul — idiome du même groupe — moins abondant, incorpore seulement l'objet de la seconde et de la troisième personne ; le magyar, plus pauvre encore, ne peut, en principe, rendre de cette façon que le pronom de troisième personne. Mais ces différentes langues ont, ce que n'a pas le basque, le verbe sans régime, le verbe indéterminé. Dans les langues sémitiques les conjugaisons « par pronoms affixes » sont, en tous cas, de véritables conjugaisons objectives.

Quant à l'incorporation des noms au verbe, habituelle, dit-on, aux langues de l'Amérique, nous n'en trouvons pas en ce moment d'exemple plus significatif que celui de l'algonquin *nadholidin* « amenez-nous le canot », formé de *naten* « amener », *amochol* « canot », *i* euphonique et *nin* « à nous » ; ou encore celui du mot chippéway *sogininjinitizoyan* « si je ne prends pas la main », dans lequel entrent *sogénât* « prendre » et *oninjina* « main ». Les formations de cette espèce ne sont qu'une simple extension du principe de l'incorporation au verbe de l'idée du régime. On a remarqué, non sans raison, qu'un certain nombre de locutions des langues romanes modernes sont de véritables exemples d'une incorporation rudimentaire ; lorsque l'italien dit *portandovi* « vous portant », *portandovela* « vous le portant », lorsque le gascon dit *deche-m droumi* « laissez-moi dormir », leur procédé nous rappelle l'incorporation du basque ou des langues américaines.

Nous pensons, en effet, avec A. Sayce, qu'il faut distinguer l'*incorporation* du *polysynthétisme*, et qu'il convient de réserver le premier de ces noms aux phénomènes que nous venons d'examiner et qui ne sont, comme on l'a vu, ni spéciaux aux langues américaines, ni assez importants pour justifier la création d'une quatrième

grande classe morphologique. A. Sayce estime même qu'il y a beaucoup plus de différence entre l'incorporation et le polysynthétisme qu'entre l'incorporation et la flexion ; pour lui, la flexion n'est, en effet, que la fusion très étroite de racines de relation avec l'élément fondamental du mot.

Nous appellerons donc *polysynthétisme* le dernier caractère signalé comme original dans les langues américaines, c'est-à-dire la composition indéfinie des mots par syncope et par ellipse. Ce caractère est le plus important.

Duponceau, qui ne confond pas l'incorporation avec le polysynthétisme, donne ce dernier comme le signe distinctif des idiomes du nouveau monde ; le même auteur assure qu'il a retrouvé ce caractère dans toutes les langues de l'Amérique à lui connues, du Groëland au Chili. Toutes réunissent un grand nombre d'idées sous la forme d'un seul et même mot. Ce mot, généralement assez long, est l'agglomération intime de mots divers, qui souvent sont réduits à de simples lettres que l'on intercale. Ainsi le groënlandais *aulisariartorasuarpok* « il s'est hâté d'aller à la pêche » est formé de *aulisar* « pêcher », *peartor* « être à faire quelque chose », *pinnesuarpok* « il se hâte », l'algonquin *pilâpé* « jeune homme non marié » est formé de *pilsitt* « chaste » et *lépâpé* « homme » *amanganachquiminchi* « chêne à larges feuilles », est formé de *amangi* « grand, gros », *nachk* « main », *quim*, terminaison des noms de fruits à coque, et *achpansi* « tronc d'arbre » ; le mexicain *notlazomahuizteopixcâtâtzin* « ô mon père, divin protecteur estimé et vénéré ! » est formé de *no* « mon », *tlazontli* « estimé », *mahuiztic* « vénéré », *teopixqui* « dieu protecteur » et *tatzi* « père » ; le chippéway *totochabo* « vin » est formé de *toto* « lait » et *chaminabo* « grappe de raisin ». Le polysynthétisme consiste donc en une

composition par syncope, tels composants perdant leurs premières syllabes, tels autres perdant leurs dernières, et il y a par conséquent entre l'incorporation et le polysynthétisme cette différence que le procédé du dernier est essentiellement syntaxique. L'incorporation remonte à la période de développement du langage, tandis que le polysynthétisme a pris naissance durant la période historique.

Ainsi le polysynthétisme n'est point un caractère primordial ; c'est une extension, ou, si l'on veut, une seconde phase de l'agglutination, et il n'y a pas de raison pour faire des langues américaines un type spécial.

Seulement, dans la série des idiomes agglutinants, ces langues viendront les dernières par ordre de progression croissante. L'on aura, par exemple, dans les premiers rangs, le dravidien avec ses formes grammaticales si peu nombreuses, puis le mandchou plus développé, puis le turc, déjà incorporant, puis les idiomes finnois dans l'ordre suivant : suomi, magyar, vogoul, mordvin, tous incorporants, puis le basque, qui est incorporant avec des tendances au polysynthétisme ; enfin les langues américaines, qui sont incorporantes et polysynthétiques. Mais cette succession, cette série ne prouvent pas plus la parenté originelle des différents idiomes dont il s'agit que certains caractères communs qu'ils peuvent avoir entre eux ne prouvent la parenté originelle des amentacées et des conifères.

Une fois entrées dans la période historique, toutes les langues d'ailleurs pourraient devenir polysynthétiques, et l'on trouve en un grand nombre d'idiomes des locutions abrégées, contractées, pleinement analogues aux syncopes des langues du nouveau monde : en latin *veneficus, stipendium, consuetudo, malo, nil, nemo, rursum*, (*venenificus, *stipi-pendium, *consueti-tudo, mavelo, magisvolo, etc.); en grec ἀμφορεύς pour *ἀμφι-φορεύς, τέτραχμον,

etc. ; en anglais *lord* pour *hlâf-weard* anglo-saxon, *world*, etc. ; *idolâtrie*, *contrôler*, *même*, *etc.* (Sayce, trad. p. 81)

Comme Duponceau l'a très bien fait remarquer, ces contractions se produisent volontiers dans les mots composés d'un usage courant, qui sont devenus peu à peu des mots à signification simple et dont on oublie la complexité originelle. C'est le basque, en Europe, qui paraît avoir le plus usé de ce procédé, et c'est pour ce motif que, dans une sériation morphologique des langues agglutinantes, on peut le placer entre les langues ouralo-altaïques et les langues américaines.

Nous ne pouvons songer à passer en revue, même de la façon la plus sommaire, les différents idiomes que nous avons énumérés ci-dessus. Nous nous bornerons (tout en faisant parfois allusion à quelques autres langues) à une esquisse générale des deux groupes les plus importants des idiomes de l'Amérique septentrionale, l'algonquin et l'iroquois. Ces deux langues d'ailleurs ne sont point parentes et elles offrent entre elles des différences notables, tant sous le rapport de la phonétique que sous celui de la formation même des mots.

L'algonquin, parlé au Canada et dans le nord des Etats-Unis, se subdivise en une trentaine de dialectes dont les principaux sont le *mikmak*, au Canada, dans la Nouvelle-Ecosse et dans les régions circonvoisines ; l'*abénaki*, dans l'Etat du Maine ; le *massachuset*, le *narraganset*, dans Rhode-Island ; le *mohican*, en Connecticut ; les idiomes de l'ancien Canada ou Nouvelle-France : l'*algonquin* proprement dit, le *chippéway* (ou *ojibway*), l'*ottawa*, le *ménoméni* et le *kri*. Les tribus iroquoises habitent la partie occidentale de l'Etat de New-York et en général la rive méridionale des grands lacs ; on peut diviser l'iroquois en *onondago*, *sénéka*, *oniéda*, *kayuga*, *tuskarora*.

La phonétique de l'algonquin est pauvre ; celle de l'iroquois est plus pauvre encore. Ils possèdent nos voyelles *a*, *e*, *i*, *o* ; quelques dialectes ont également *u*. Ils connaissent les deux demi-voyelles *y* et *w*, dont la seconde se change en une espèce de sifflante labiale ; c'est ce son que les missionnaires ont transcrit et transcrivent encore par le chiffre 8, sous prétexte que ce signe ressemble à la ligature « ou » des Grecs, et que le mot « huit » rappelle sa prononciation.

L'algonquin possède les deux gutturales *k*, *g*, tandis que l'iroquois n'a qu'un son de cette sorte, transcrit parfois *g*, le plus souvent *k*. Les deux idiomes connaissent la palatale *tch*, certains dialectes algonquins ont le *dj* ; l'algonquin emploie *t* et *d*, l'iroquois *t* seulement ; dans ce dernier idiome, point de labiales : l'algonquin possède *p* et *b*. Chacune de ces langues connaît les nasales appartenant au même ordre que ses explosives. L'une et l'autre ont *r* et *l*, toujours permutants, souvent indistincts. En algonquin, les sifflantes sont assez nombreuses: on y retrouve *h*, *ch* dur allemand, *s*, *z* et *j* français ; mais en iroquois l'on ne cite que *h* et *s* (*f* est dialectal) ; ajoutons que ces idiomes ont au moins trois voyelles nasales, *an*, *en*, *on*. Le seul son qui offre quelque difficulté aux Européens serait le *w* placé devant une consonne. Duponceau dit à ce sujet : « C'est comme *ou* dans notre mot *oui*, mais suivi immédiatement d'une consonne et prononcé sans repos intermédiaire, ce qui l'a fait appeler l'*ou* ou le *w* sifflé, parce qu'en effet il faut siffler pour le prononcer. Le même son existe dans l'abénaki ; mais, au lieu d'être labial, comme dans le lénapé, il est guttural et se prononce du fond de la gorge. On ne le trouve point dans l'algonquin proprement dit, ni dans le chippéway. Il n'est pas non plus dans la langue des Ottawais, ils y substituent l'*ou* voyelle. Ainsi, tandis

qu'un Lénapé prononcera *w'danis* « sa fille » (en sifflant le *w*), l'Ottawais dira *oudanis.* »

Toujours suivant Duponceau, les Indiens Algonquins articulent d'une façon très distincte et prononcent les voyelles très ouvertes : les brèves ont l'accent « frappé » et les longues l'accent « appuyé » ; la dernière syllabe de la phrase est articulée avec une grande énergie. La prononciation des Américains du Sud est plus rude que celle des Américains du Nord.

L'article, que certains auteurs ne veulent pas reconnaître, est retrouvé par Duponceau au moins en algonquin. C'est, comme d'ordinaire, un pronom démonstratif, *monko* (en massachuset), réduit à *m* préfixé ; mais aujourd'hui son existence est méconnue au point qu'on le conserve concurremment avec les affixes possessifs. Ainsi le chippéway dit *mittig* « arbre » et *ki mittig* « ton arbre » ; tandis que le lépané dit *kittuk* « arbre » *m'hittuk* « l'arbre » et *k'hittuk* « son arbre ». L'article se rencontre encore dans d'autres langues américaines : en iroquois sous la forme *ne*, en otomi sous la forme *na*. Ce qui l'a fait méconnaître, c'est la tendance que possèdent ces langues à la spécialisation, à la détermination, et qui fait que, chez elles, les noms sont toujours accompagnés d'un affixe possessif.

En algonquin, l'on ignore la distinction des genres ; l'on en compte deux en iroquois : les grammairiens leur donnent le nom de genre « noble » et de genre « ignoble ». Le premier s'applique aux divinités et à la partie mâle du genre humain ; le second embrasse tout ce qui n'est pas du premier genre. Il y a des particules ou des affixes différents pour les êtres animés et pour ceux qui ne le sont point.

La conjugaison nominale — ou, pour mieux dire, et ainsi que nous l'avons expliqué plus haut, la dérivation possessive — se forme par l'adjonction des éléments pro-

nominaux en tête du nom. L'adjectif, toujours invariable, se place de même, en algonquin, avant le mot qualifié ; c'est ainsi que *kuligatchis* « ta jolie petite patte » est formé de *ki* « toi » *wulit* « jolie », *wichgatt* « patte » et du diminutif *chis ; kitanittowit* « le grand esprit » est formé de *kita* « grand », *manitu* « esprit » et de la terminaison adjective *wit.*

Le verbe algonquin peut être « absolu », c'est-à-dire sans régime ; « transitif », c'est-à-dire pourvu d'un régime direct, et « passif », avec un régime indirect. On s'est plu, mais à tort, à reconnaître chez lui un grand nombre de modes : il n'en possède aucun en réalité ; il se forme seulement un conditionnel en intercalant une particule. Le verbe iroquois est également absolu, réfléchi, réciproque, passif et transitif, à régime direct et indirect. De plus, il y aurait dans certaines langues américaines des traces d'une soi-disant conjugaison pour ainsi dire sexuelle ; ainsi, en abénaki, un homme dirait *nénananbasanbai* « je n'ai pas beaucoup d'esprit » ; une femme dirait *nénananbaseskouai.* L'on commence à comprendre que, grâce à ces nombreuses variations, un missionnaire anglais Edwin James, ait pu attribuer au verbe chippéway de six à huit mille formes.

Pas plus que dans les langues dravidiennes, on ne peut exprimer en algonquin ou en iroquois les verbes « être » et « avoir » dans leur sens réputé absolu. C'est ainsi que cette phrase : « je suis un homme », se dira en narraganset *ninin* « moi homme », en lénapé *lenno n'hackey* « un homme mon corps » ; pour dire « à qui est ce canot ? » l'on demandera en ottowa *watchimânet* « à qui canot ? » ; en ménoméni *wahotosoyâwik* « qui possède canot ? »

En somme, le vocabulaire de ces idiomes est assez pauvre. Il leur manque, comme bien l'on pense, la plupart des mots capables d'exprimer une idée abstraite, et qu'ils

remplacent soit par des mots empruntés à l'anglais, à l'espagnol, au français et même à l'allemand, soit par des périphrases développées que les grammairiens aiment souvent à appeler des mots de dix à douze syllabes. Dans les dialectes algonquins, les cinq premiers nombres sont des mots simples et paraissent seuls primitifs : « dix » serait « cinq de plus (que cinq) », « cent » serait « dix fois dix » et « mille » « la grande dizaine de dizaines » ; l'iroquois, au contraire, semble avoir compté jusqu'à dix. Il y aurait de fort curieuses remarques à faire sur les noms de parenté, très abondants, en iroquois par exemple. On les a classés par catégories : parenté supérieure : père, mère ; — inférieure : fils, frère, cadet ; — affinité supérieure : beau-père ; — affinité inférieure : bru ; — parenté corrélative : beau-frère, etc. Les langues dravidiennes sont particulièrement riches en mots de cette nature : on y distingue, par exemple, les frères aînés des frères cadets par des termes différents, de même que l'on distingue en basque la sœur d'une femme de la sœur d'un homme. La raison de ces complications est sans nul doute, nous semble-t-il, dans le manque d'expressions générales qui est habituel aux idiomes inférieurs,et les philologues ou les géographes n'ont que trop souvent admiré cette richesse apparente de mots.

Nous ne pourrions, sans dépasser les bornes qui s'imposent au présent volume, examiner la phonétique et les éléments de formation propres aux différents groupes des langues américaines. Nous eussions désiré citer plus d'exemples que nous ne l'avons fait, reproduire quelques phrases entières et les analyser. Les langues américaines ont donné lieu, et donnent lieu encore, à des travaux si dépourvus de méthode, qu'on ne saurait trop s'attacher à faire voir comment elles se relient aux autres idiomes agglutinants.

Le point important était de faire bien saisir ce qu'il

faut entendre par ces expressions de *polysynthétisme*, et d'*incorporation*. Si nous avons réussi à exposer nettement ce qui en est, nous pensons avoir dit des langues américaines ce qui est essentiel.

Disons en terminant que la numération est vigésimale dans certaines langues américaines, particulièrement au centre, décimale dans certaines autres langues (Algonquins, Iroquois, Californiens, Araucans, etc.).

§ 17. Les langues hyperboréennes.

On comprend sous ce nom géographique l'ensemble des langues parlées dans les régions arctiques.

Le *youkaghir* est parlé par un millier d'individus au nord-est de la Sibérie, immédiatement à l'est du yakout (langue du groupe turc).

Le *tchouktche* asiatique et le *koriaque* sont parlés plus à l'est encore, tout au nord-est de la Sibérie ; ces deux idiomes sont proches parents.

Au sud de la presqu'île de Kamtchatka est parlé le *kamtchadal*.

Plus au sud encore, dans les îles Kouréliennes et les îles japonaises septentrionales, nous trouvons la langue des *Aïnos*.

Celle des *Ghiliaks* appartient au continent.

Au centre même de la Sibérie nous trouvons le *iénisséin* (ou ostiaque iénisséin) et le *kotte*.

Les dialectes *innuits* sont parlés par les Esquimaux tout au nord de l'Amérique. Le *tchouktche* américain, qu'il ne faut pas confondre avec le tchouktche asiatique, est parent des dialectes innuits ; il est parlé sur la côte nord-occidentale.

Les dialectes *aléoutiens* sont essentiellement distincts des dialectes innuits.

Au surplus, le nom d'*hyperboréennes* ou *arctiques*,

sous lequel on réunit ces différentes langues, ne doit pas donner le change sur leur plus ou moins d'affinité soit entre elles, soit avec d'autres idiomes.

Bien des hypothèses sont encore permises à ce sujet, mais il est vraisemblable qu'un certain nombre de ces idiomes résisteront à toutes les tentatives que l'on pourra faire en vue de les laisser parmi tel ou tel groupe suffisamment connu.

Il serait dangereux, en tout cas, d'accorder aux relations des missioniaires sur telle ou telle de ces langues, notamment sur celles des Esquimaux, plus de crédit qu'il ne convient. On n'y trouve, le plus souvent, que des rapprochements de mots, des étymologies ; en somme, rien de scientifique. Ajoutons d'autre part que certains idiomes hyperboréens ont été étudiés avec soin par des auteurs compétents, ainsi qu'on peut le voir dans les publications de l'Académie de Pétersbourg.

Il y a à remarquer dans la langue des Youkaghirs que le nombre est rendu dans le pronom par une véritable flexion : *mot* « je », *mit* « nous ». Celle des Aïnos peut être citée comme une langue agglutinante sortant à peine de la phase monosyllabique : point de distinction formelle du nombre dans les noms, en principe du moins ; les cas nominatif et accusatif ne sont indiqués par aucun élément accessoire ; dans le verbe, le temps, le mode, la personne, sont exprimés par des procédés périphrastiques ; la numération est vigésimale. La langue des Aléoutes est par contre à une période développée de l'agglutination ; le système de numération est quinaire.

§ 18. Les langues du Caucase.

Les langues du Caucase (ainsi appelées de la région où elles sont parlées) se divisent en deux groupes principaux, celui du nord, celui du sud.

Parlons d'abord du GROUPE SEPTENTRIONAL.

Longeant le versant nord du Caucase, de la mer Caspienne à l'est, jusqu'au détroit de Kertch (entre la mer d'Asov et la mer Noire) à l'ouest, il comprend à son tour trois groupes distincts : le groupe *lesghien*, à l'est (dans le Daghestan), qui confine à la mer Caspienne et semble compter quatre cent mille individus ; le groupe *kiste*, au centre, beaucoup moins considérable que le précédent ; le groupe *tcherkesse* ou circassien, à l'ouest, qui occupe presque toute la moitié nord-occidentale du Caucase et n'est peut-être pas éloigné de comprendre à lui seul tout autant d'individus que n'en comprennent les deux groupes précédents.

Un mot maintenant sur chacun de ces trois sous-groupes et leurs propres subdivisions.

Le lesghien comprend, entres autres idiomes, l'*avare* (1), *le kasi-koumuque* (ou lak), l'*akoucha*, le *kurine*, l'*oude*.

Le kiste (ou tchetchenze), à l'ouest du Daghestan, ainsi que nous l'avons dit, comprend l'*ingouche* (ou lamour), le *karaboulak*, le *tchetchenze* proprement dit, le *thouche* (ou mosok), qui, bien qu'appartenant au groupe septentrional, est parlé au sud du Caucase, vers la source de l'Alasan. On porte à cent quarante mille le nombre des individus parlant les dialectes kistes.

Il est assez difficile d'évaluer le nombre actuel des Circassiens du Caucase : on le portait autrefois au chiffre de quatre cent quatre-vingt-quinze mille, mais beaucoup d'entre eux ont été attirés récemment dans la Turquie européenne.

Le GROUPE MÉRIDIONAL est formé du géorgien, du suane, du mingrélien, du laze.

Le *géorgien* serait parlé par près de trois cent mille individus.

(1) SMIRNOW, *Notice sur les Avares du Daghestan*. Revue d'Anthropologie, t. V, p. 84. Paris, 1876.

Le territoire du *suane* est situé au nord-ouest du géorgien.

Au sud du suane et à l'ouest du géorgien est parlé le *mingrélien.*

Le *laze*, enfin, est parlé plus au sud encore, au sud-est de la mer Noire, dans le Lazistan, pays soumis aux Ottomans.

Ces quatre derniers idiomes remonteraient à une origine commune, mais il est loin d'être établi qu'ils soient apparentés aux idiomes du versant septentrional, et même que ceux-ci soient alliés entre eux. Un certain nombre des langues du Caucase ont été cependant étudiées de très près, notamment par Schiefner dans les *Mémoires de l'Académie de Pétersbourg.*

Le matériel phonique de tous ces idiomes est riche lorsqu'il s'agit des consonnes, assez pauvre lorsqu'il s'agit des voyelles.

On trouve dans ces langues la dérivation par suffixes, celle par préfixes, celle par infixes. Les cas sont ordinairement formés au moyen de suffixes : laze, *bozo* « fille », *bozos* « à la fille », *bozote* « avec la fille » ; de même le pluriel : thouche *dad* « père », *khok* « pied », pluriel *dadi, khoki.* Dans ses formes verbales l'abkhase préfixe l'élément personnel dérivatif : *si-qoup* « je suis », *u-qoup* « tu es », *ha-qoup* « nous sommes ». Les idiomes méridionaux forment les noms d'agents au moyen du préfixe *me : puri* « pain », *mepure* « fabricant de pain » ; *thewzi* « poisson », *methewze* « pêcheur ». L'abkhase infixe dans la racine l'élément pronominal.

Certaines langues du Caucase ont un système décimal de numération ; chez la plupart le système est vigésimal.

L'on a tenté souvent d'identifier les langues dites *du Caucase* avec les langues indo-européennes ou les langues sémitiques, mais cela a toujours été sans succès. Nous pensons qu'il faut les regarder comme complète-

ment distinctes des autres groupes linguistiques, même du groupe ouralo-altaïque.

§ 19. De quelques autres idiomes classés parmi les langues agglutinantes.

Dans le nombre des langues que nous venons de passer en revue, il y en a plusieurs, sans doute, qui demanderaient à être étudiées de plus près qu'elles ne l'ont été jusqu'à présent, et qui, en réalité, ne sont connues que d'une façon très imparfaite.

On est moins bien renseigné encore sur quelques autres idiomes, les uns actuellement vivants, les autres disparus.

A. *La prétendue langue scythique.*

Cette expression de « cythique » a été employée de deux façons différentes ; on l'a appliquée soit à un peuple particulier, soit à un ensemble de populations plus ou moins apparentées. Dans le premier cas, l'on ne suppose qu'une seule langue scythique, un seul peuple scythe ; dans le second l'on suppose, non plus une race et une langue scythique, mais bien des races et des langues scythiques. La première opinion a été peu soutenue ; la seconde, par contre, n'a pas laissé de séduire des auteurs aussi compétents que l'est, par exemple, D. Withney. Celui-ci, en effet, donne aux langues ouralo-altaïques ce nom de langues « scythiques » (1), nom que les Grecs donnaient aux

(1) *Language and the Study of Language*, 3e édit., p. 309. Londres, 1879.

races nomades du Nord-Est, sinon à toutes ces races, du moins à plusieurs d'entre elles.

Nous ne pouvons admettre cette appellation, beaucoup trop étendue, nous semble-t-il. Il est vraisemblable, sans doute, bien que l'on ne puisse en fournir une preuve directe (1), que l'antiquité a classé parmi ses Scythes plus d'une peuplade appartenant au groupe ouralo-altaïque ; mais, d'autre part, il semble avéré qu'elle a également donné le nom de Scythes à des peuples parlant un idiome indo-européen. La langue des Scythes du Pont paraît avoir été une langue éranienne, ainsi qu'a cherché à le démontrer Müllenhoff.

Quelques auteurs ont pu penser sans trop d'invraisemblance qu'une partie des Scythes parlaient un idiome se rattachant aux langues slaves (2). Nous estimons, en somme (3), avec Frédéric Müller, que ce nom général de Scythes n'est qu'une expression géographique ne répondant à aucune idée précise de race ou de langue. La Scythie est simplement le nord de l'Europe et de l'Asie, les Scythes sont les habitants de cette région.

Ainsi il nous semble très peu prudent, pour le moins, de parler d'une langue scythique, et de donner ce nom à l'ensemble des idiomes ouralo-altaïques.

B. *La langue de la seconde colonne des inscriptions cunéiformes.*

La première colonne des inscriptions trilingues des Achéménides est, comme l'on sait, rédigée en vieux

(1) Schiefner, *Sprachliche Bedenken gegen das Mongolenthum der Skythen*. Mélanges asiatiques, t. II, p. 531. 1856.

(2) Consultez Gr. Krek, *Einleitung in die slavische Literaturgeschichte und darstellung ihrer ælteron period*, t. I, p. 36. Gratz, 1874. Consultez également Spiegel, *Eränische Alterthumskunde*, t. I, p. 333 et suiv. Leipzig, 1873.

(3) Cf. Fligier, *Zur Scythenfrage*, 1878.

perse ; c'est celle qui fut déchiffrée la première. La troisième colonne, dont la langue ne fut connue que longtemps après, est en assyrien, dialecte sémitique.

On a donné différents noms à la langue de la seconde colonne de ces inscriptions, entre autres celui de langue médique et celui de langue scythique. Ce dernier nom, qu'ont proposé, et dont se sont servis H. Rawlinson (1). et Norris (2), est beaucoup trop vague, ainsi que nous l'avons dit au paragraphe précédent, pour convenir à un idiome quelconque. Celui de médique paraît avoir plus de raison d'être. On allègue en sa faveur que certaines inscriptions rédigées dans la langue de la deuxième colonne des monuments achéménides, ont été trouvées dans les régions de l'ancienne Médie, sans être accompagnées de versions éranienne ou assyrienne. Les trois langues des inscriptions cunéiformes devaient, ajoute-t-on, être les trois langues des principaux peuples de l'empire ; or, la première étant perse, la troisième assyrienne, la seconde ne pouvait être que médique (3).

Pour E. Norris, le soi-disant médique était un idiome du groupe ouralo-altaïque, proche parent du magyar, de l'ostiaque, du permien et des autres langues de la même famille. A. Mordtmann en fit également un idiome ouralo-altaïque, mais il le rattacha au groupe turc ou tatar (4), supposa qu'un certain nombre d'éléments indo-européens s'y étaient glissés à différentes époques, et lui donna le nom de susien, langue de la Susiane.

(1) *Notes on the early Historia of Babylonia.* Journal of the Royal Asiatic Society, t. XV, p. 215.

(2) *Mémoir on the Scythic Version of the Behistun Inscription.* Journal of the Royal Asiatic Society, t. XV, p. 1. Londres, 1863.

(3) Benfey, *Geschichte der Sprachwissenschaft und orientalischen Philologie in Deutschland*, p. 633. Munich, 1869.

(4) *Ueber die Keilinschriften zweiter gattung.* Zeitschrift der deutschen morgenlændischen Gesellschaft, t. XXIV, p. 76. Leipzig, 1870.

J. Oppert, lui aussi, s'est occupé de cette question (1). Après avoir adopté le nom de « scythique », il s'est rallié à celui de médique. La langue de la seconde colonne des inscriptions cunéiformes serait la langue de la dynastie médique, la langue d'Arbacès, de Déjocès, de Cyaxare et autres, dont les noms, sous cette forme, sont simplement rendus à la façon des Perses. Cette dynastie aurait régné de l'année 788 à l'année 560 avant notre ère, deux siècles et un quart de siècle ; sa langue et sa religion auraient été tout à fait distinctes de la langue et de la religion des Achéménides. (Voir Oppert, *Congrès des Orientalistes*, tome II, p. 181 ; *Zeitschrift der morgenlœndischen Gesellschaft*, tome III, p. 1) ; *le Peuple et la Langue des Mèdes*, 1879).

En fin de compte, la question porte principalement sur ces deux points : la langue de la seconde colonne des inscriptions cunéiformes appartient-elle au groupe des langues ouralo-altaïques ? Cette langue est-elle la langue des Mèdes ?

Sur le premier point, nous pouvons répondre avec F. Spiegel (2) que la langue dont il s'agit ne nous paraît pas encore suffisamment déchiffrée. Les auteurs que nous avons cités plus haut (et auxquels on peut en joindre quelques autres, par exemple N. Westergaard) sont loin d'avoir fait partager à tous les juges compétents leur opinion sur le caractère finnois ou tatar de l'idiome dont il s'agit. Caldwell veut le rapprocher des langues dravidiennes de l'Inde méridionale ; cette manière de voir est inacceptable. En l'état actuel de la question, il semble prudent d'attendre les résultats de nouvelles recherches.

D'autre part, nous devons nous demander s'il n'est pas téméraire de regarder définitivement les Mèdes comme

(1) *On the Median Dynasty. Its Nationality and its Chronology*, 1874.
(2) *Eränische Alterthumskunde*, t. I, p. 381. Leipzig, 1871.

un peuple d'origine ouralo-altaïque. F. Spiegel ne peut se ranger à cette opinion et ses arguments ne nous paraissent point sans valeur. Le témoignage d'Hérodote est formel, celui de Strabon ne l'est pas moins. L'un et l'autre tiennent les Mèdes pour un peuple aryen. Leurs noms propres, leurs noms géographiques s'expliquent tous, d'ailleurs, par les langues éraniennes et non par le suomi ou par le turc (1).

Jusqu'à plus ample information, il paraît donc raisonnable de ne classer encore dans aucun groupe, dans aucune famille même, la langue de la seconde colonne des inscriptions cunéiformes et de ne lui donner aucun nom particulier.

C. *La langue dite sumérienne ou accadienne.*

On a supposé, il n'y a pas longues années, que des populations parlant une langue agglutinante avaient occupé avant les Assyriens les régions de la Babylonie, et que, pour s'implanter dans le pays, la civilisation sémitique avait dû se greffer sur cette civilisation antérieure. Hincks donna à cette langue le nom biblique d' « accadien », qui semble jouir aujourd'hui d'une certaine faveur, mais que son auteur, paraît-il, ne proposait que sous toutes réserves. Pour Oppert, le nom d' « accadien » n'est qu'un synonyme d'assyrien ; assyrien et accadien seraient la langue sémitique de Ninive et de Babylone, la langue de la troisième colonne des inscriptions cunéiformes. Au peuple qui avait précédé les Sémites en Assyrie (et leur aurait communiqué son écriture cunéiforme et sa civilisation) il donna le nom de « Kasdo-scythique » ou sumérien ; à la langue de ce peuple il donna le nom de « sumérien » (2). Rawlinson nomma cet idiome le proto-chaldéen.

(1) Spiegel, *Op. cit.*, t. I, p. 384.
(2) *Congrès des Orientalistes*, session de Paris, 1876, t. II, p. 134.

La langue sumérienne aurait disparu, à un moment donné, de l'usage populaire, mais les prêtres l'auraient conservée dans l'exercice du culte. Oppert a tenté le premier d'esquisser les traits essentiels d'une grammaire sumérienne (1).

On a cherché depuis lors avec beaucoup d'ardeur à restituer le lexique sumérien et à le comparer à celui des langues ouralo-altaïques. On professa que les prédécesseurs des Sémites assyriens sur le sol de la Babylonie avaient parlé une langue alliée au finnois ; qu'ils avaient connu une grande civilisation ; qu'ils avaient communiqué aux Assyriens immigrant au milieu d'eux leur système d'écriture cunéiforme ; enfin, qu'avant de perdre leur propre langue, ils avaient initié les nouveaux venus à une culture à laquelle ces derniers étaient restés étrangers jusque-là.

La thèse du sumérien ne s'imposa point de haute lutte. Joseph Halévy a tenté une explication tout autre des textes soi-disant accadiens. Il a cherché d'abord à démontrer que la langue en question n'a rien de commun avec les idiomes ouralo-altaïques : sa phonétique en diffère largement ; ses racines n'ont ni la forme, ni l'usage de celles des langues ouralo-altaïques ; le mode de structure des mots est tout différent ; les pronoms n'ont rien de commun ; la conjugaison est construite en de tout autres conditions ; les deux vocabulaires, enfin, ne sauraient être sérieusement comparés : c'est à peine si l'on rapproche entre eux quelques dizaines de mots dits accadiens, d'un nombre égal de mots empruntés aux divers idiomes du groupe finnois. En somme, l'existence d'un peuple parlant une langue ouralo-altaïque sur le sol de la Mésopotamie, ne serait démontrée ni par les monuments (qui appartiennent tous à l'art sémitique), ni par

(1) *Journal asiatique*, VIIe série, t. I, p. 113. Paris, 1873. *Revue de philol. et d'ethnogr.*, 1875.

les noms géographiques, également sémitiques, ni par le témoignage des auteurs. Les textes accadiens ne seraient, en définitive, que de l'assyrien pur et simple, écrit, non plus à l'aide d'un système phonétique, mais bien à l'aide de monogrammes artificiellement combinés.

En autres termes, des deux côtés on aurait de l'assyrien, uniquement de l'assyrien, mais les textes soi-disant sumériens seraient écrits en idéogrammes au lieu de l'être en caractères phonétiques (1) ; ce qu'on appelle accadien ou sumérien serait simplement une façon d'écrire l'assyrien, une écriture hiératique, sorte de transition entre le système idéographique et le système syllabique.

Stanislas Guyard, lui aussi, n'a voulu voir dans le prétendu suméro-accadien que des mots assyriens écrits dans un système particulier (2).

Renan s'est exprimé sur cette question avec une grande réserve. Avant l'arrivée des Assyriens, avant l'arrivée des Eraniens, il y avait en Babylonie une véritable civilisation. Il est très probable que cette civilisation a possédé en propre, a créé l'écriture dite « cunéiforme » ; mais faire de la langue de ces prédécesseurs des Assyriens un idiome ouralo-altaïque, cela dépasse toute permission légitime. Il y a lieu d'être étonné de voir rattacher « cette antique substruction de la civilisation savante de Babylone aux races turques, finnoises, hongroises, à des races qui n'ont guère su que détruire et qui ne se sont jamais créé une civilisation propre. Le vrai peut quelquefois ne pas être vraisemblable, et si l'on nous prouve que ce sont des Turcs, des Finnois, des Hongrois qui ont fondé

(1) *Revue de philol. et d'ethnogr.*, t. III, *Revue critique*, 14 juillet 1884.

(2) *Journal asiat.*, 1880, t. I, p. 540. — *Revue de l'histoire des religions*, 1880-1882. — *Revue critique*, 17 juillet 1882. — *Zeitschr, für Keilschriftforschung*, 1884. — Consultez : Sayce, *Princ. de philologie ; Schrader, Zeitschr, der deutschen morgenl. Gesellsch.*, t. IX ; le même recueil, passim ; Haupt (P.), *Die sumerischen Familiengesetze*, 1879.

la plus puissante et la plus intelligente des civilisations anté-sémitiques et anté-aryennes, nous croirons ; toute considération *a priori* doit être subordonnée aux preuves *a posteriori*. Mais la force de ces preuves doit être en proportion de ce que le résultat a d'improbable. »

Fr. Müller dit, de son côté : « L'écriture cunéiforme est due selon toute vraisemblance à un peuple dont la position ethnologique est inconnue, le peuple d'Accad. On a voulu en faire un peuple touranien, ou, pour parler en termes plus précis, ouralo-altaïque, aliié particulièrement au rameau finnois. Abstraction faite du défaut de méthode qui a conduit à cette opinion, nous la regardons comme inconciliable avec ce que nous savons de l'ethnologie de la haute Asie (1). »

Quoi que l'on en dise chaque jour, les preuves de la théorie sumérienne ne sont pas définitives. Nous nous gardons bien toutefois de combattre et de nier le sumérien ; nous sommes tout disposé à l'accepter, à le classer au nombre des langues agglutinantes, à le rapprocher même du groupe finnois ; mais nous en sommes encore à attendre des arguments décisifs, une grammaire véritable. Nous ne pouvons nous contenter des centaines d'étymologies plus ou moins ingénieuses que l'on a accumulées les unes sur les autres avec autant d'intrépidité que peu de critique. Ce procédé est bien connu ; n'a-t-il pas triomphé des inscriptions étrusques, des inscriptions gauloises, des inscriptions ibériennes ?

On écrit beaucoup sur l'accadien, beaucoup trop peut-être. Il ne suffirait, pour le faire accepter, que d'un court travail fait avec méthode. Il se peut que la démonstration décisive ne tarde pas à venir, mais jusqu'à ce jour, nous le répétons, elle n'a pas été fournie. Les défenseurs de

(1) *Grundriss der Sprachwissenschaft*, t. I, p. 168. Vienne, 1876.

la théorie sumérienne doivent, avant tout, être parfaitement au courant de la phonétique, de la structure et du vocabulaire spécial des langues ouralo-altaïques. Tous les auteurs qui ont écrit sur le sumérien ne paraissent pas posséder à fond cette première condition de leur comparaison. Le cas d'un certain nombre d'entre eux semble être plus grave encore. Il s'agirait, paraît-il, d'apporter un nouvel appui à la vieille tentative de l'accommodement des origines ethniques et linguistiques au meilleur gré de la fable hébraïco-chrétienne.

Il importe que la question ne soit pas étendue en dehors des limites qui lui conviennent. La seule question engagée est celle du caractère agglutinatif des textes dits « accadiens » ou sumériens ; on ne saurait la compliquer sans risquer d'entraver, par un manque regrettable de bon vouloir, le cours d'une enquête qui doit être purement linguistique.

§ 20. La théorie des langues touraniennes.

Dans la période de formation des sciences nouvelles, alors que l'on songe avant tout à grouper et à classer les premiers résultats acquis, il se glisse souvent de ces théories générales qui peuvent réduire les esprits amateurs de choses simples et faciles, mais qui sont appelées à s'écrouler piteusement, un jour ou l'autre, sous les entreprises de la critique.

La linguistique n'a pas échappé aux théories de cette espèce.

On peut ranger parmi les conceptions les plus fantaisistes la théorie d'une famille *touranienne*, qui, malgré son invraisemblance, n'a pas laissé de jouir jusqu'en ces derniers temps d'un certain crédit.

Hâtons-nous de le dire, cette théorie ne repose sur au-

cun fait scientifique, et elle n'a été imaginée que pour soutenir des conceptions ethnographiques très peu sérieuses.

Tout à l'heure, nous nous occuperons de son origine et de son nom ; cherchons, auparavant, à établir en quoi elle consiste.

Et d'abord il faut distinguer. Il existe deux variétés de « touranisants » ; le touranisant absolu et le touranisant modéré.

Celui de la première variété est le touranisant orthodoxe. Pour lui, toutes les langues étrangères au groupe européen et aux groupes sémitique et khamitique constituent un groupe « touranien ». Les idiomes de ce groupe auraient en commun, non seulement un ensemble de procédés de structure, mais encore un grand nombre de racines : il y aurait donc eu une langue commune, une langue mère touranienne. On admet d'ailleurs dans ce groupe une division septentrionale et une division méridionale. La première comprendrait les langues ouralo-altaïques, dont nous avons parlé dans le paragraphe 14 de ce chapitre ; la seconde comprendrait, non seulement toutes les autres langues agglutinantes, mais encore les langues monosyllabiques de l'extrême Orient.

La seconde variété est celle du touranisme dissident, du touranisme hétérodoxe. Et ici encore il faut distinguer.

Une première sous-variété ne croit déjà plus au touranisme, mais elle cherche à en sauver au moins le nom. Ces touranisants de troisième degré donnent le nom de touraniennes aux langues que nous appelons « ouralo-altaïques », ou simplement altaïques, et qui se divisent, comme nous l'avons vu, en cinq groupes : samoyède, finnois, turc, mongol, tongouse.

Moins osée que la précédente, la seconde sous-variété des touranisants modérés compose le groupe touranien,

non seulement des idiomes ouralo-altaïques, mais encore des langues dravidiennes, des langues maléo-polynésiennes, du tibétain, puis du siamois. Nous exposons et ne critiquons pas.

Nous ne demandons point, par exemple, pourquoi l'on fait abstraction ici du chinois, du japonais, du hottentot et des autres idiomes que les touranisants absolus appellent également touraniens.

Cette théorie, avons-nous dit, a été imaginée en grande partie pour venir à l'aide de données ethnographiques non moins imaginaires, mais par contre très hortodoxes. C'est une théorie bonne à abuser les personnes crédules ou celles qui n'ont ni le temps ni la faculté de contrôler par elles-mêmes les assertions qu'on leur propose au nom même de la science. Un patriarche du nom de Tour aurait donné naissance à une race touranienne dont la langue aurait été la mère commune des différents idiomes soi-disant touraniens. Une légende persane fut habilement greffée sur cette invention, et l'orthodoxie judaïco-chrétienne battit monnaie de cette nouvelle théorie, qui, pour être dépourvue de tout caractère sérieux, ne lui semblait pas moins bonne à recueillir, puisqu'elle s'accommodait sans peine à l'enseignement de ses livres saints.

S'il est un fait avéré, c'est bien celui qu'ont démontré Schleicher, W. Withney, et tant d'autres avec eux, que ces prétendues langues touraniennes n'ont de commun qu'une chose : le nom ridicule qu'on se plaît à leur donner. La structure générale du basque, du japonais, du magyar est sans doute la même ; toutes ces langues, sans doute, suffixent aux noms des éléments dont le rôle est parfaitement analogue ; toutes ces langues, en un mot, sont de forme agglutinante, mais les éléments qui constituent le fond même de chacune d'elles sont tout différents, leurs racines sont irréductibles.

Il ne suffit point, pour proclamer audacieusement

qu'elles remontent à une seule et même source, de n'avoir pu, ni de près ni de loin, les ramener à une forme commune.

Ce qui distingue avant tout les touranisants, c'est leur grand aplomb. Il ne faut cependant qu'un peu de critique pour les ruiner entièrement.

Il est fâcheux, en tous cas, que certains auteurs fassent à ce nom fantaisiste de langues touraniennes l'honneur de le regarder — tout en le condamnant — comme un fait dont il n'y a plus à se débarrasser. C'est pour cette condescendance même que l'on arriverait à lui donner encore quelques beaux jours, sinon à l'implanter tout à fait. Le meilleur moyen de combattre la théorie touranienne est peut-être de n'en plus parler.

Le nom malencontreux de langues sémitiques répond, au moins, à un ensemble de choses bien défini, et on peut l'accepter sous toutes réserves ; mais celui de touranien et de langues touraniennes n'est fait que pour perpétuer les plus graves erreurs.

CHAPITRE V.

TROISIÈME FORME LINGUISTIQUE : LA FLEXION.

Nous arrivons maintenant à la troisième forme du langage : la flexion.

Nous avons vu que dans la période du monosyllabisme la racine et le mot étaient tout un, que la phrase n'était qu'une succession de racines monosyllabiques, isolées les unes des autres. Dans la seconde période nous avons vu que certaines racines, passant de la condition de mots indépendants à l'état de simples suffixes (par exemple dans les langues ouralo-altaïques) ou de simples préfixes (dans les idiomes du système bantou), ne servent plus qu'à exprimer les relations, actives ou passives, des racines qui ont conservé leur pleine et entière signification.

Dans la première période la formule du mot, ainsi que nous l'avons dit, est simplement R et la formule de la phrase est R+R+R, etc. ; par R nous entendons la racine.

Si nous représentons par r les racines dont le sens primitif s'est oblitéré et qui ne servent plus de préfixes ou de suffixes, nous avons comme formules de mots, dans la seconde période, Rr, Rrr, rR, rRr et nombre de combinaisons analogues.

Deux systèmes de langues, celui des langues sémito-khamitiques et celui des langues indo-européennes, après avoir connu la période du monosyllabisme, puis celle de l'agglutination, arrivèrent indépendamment l'une de l'autre à la troisième phase, celle de la flexion.

§ 1. Quest-ce que la flexion ?

Ici la racine peut exprimer, par une modification de sa propre forme, les rapports qu'elle a avec telle ou telle autre racine. La flexion, c'est la possibilité pour une racine d'exprimer, en se modifiant ainsi, une certaine modification du sens.

Dans tous les mots d'une langue à flexion, la racine *n'est pas nécessairement modifiée*, elle demeure parfois telle quelle, comme dans la période d'agglutination, mais *elle peut être modifiée*.

Les langues dans lesquelles les relations que les mots affectent entre eux peuvent être ainsi exprimés, non seulement par l'annexion de suffixes et de préfixes, mais encore par une variation de la forme même de la racine, sont des langues à flexion. Certains auteurs leur donnent le nom de langues inflectives.

Si nous représentons par un exposant x cette puissance de la racine, la formule RR de l'agglutination peut devenir RXR dans la période de la flexion, la formule RR peut devenir RRX, la formule RRR peut devenir RRXR, et ainsi de suite.

Il y a plus. Non seulement la racine que les Chinois auraient appelée « pleine » peut recevoir cet exposant, comme nous le voyons dans la formule précédente, mais la racine qui forme l'élément de relation, le suffixe, peut également être modifiée.

Dans une langue à flexion, la formule d'un mot peut donc être RRX, RXRX, RRRX et ainsi de suite, c'est-à-dire qu'un élément dérivatif peut être modifié dans sa forme, en vue d'un changement de sens, tout comme peut l'être la racine principale elle-même.

« Si nous acceptons, dit Fr. Müller (*Grundriss*, t. I, p. 136), que les langues ont eu pour origine de simples

racines — axiome généralement admis — le développement linguistique n'a pu se faire que par composition de racines les unes avec les autres. D'abord, les racines se sont simplement juxtaposées, puis la fusion est devenue plus intime. » Cette conception diffère, on le voit, de celle que nous venons d'exposer ; pour Fr. Müller, la différence entre l'agglutination et la flexion consiste en ce que la flexion est une agglutination moins reconnaissable au premier coup d'œil.

Victor Henry (1) explique le passage de la phase agglutinante à la phase de flexion, dans les langues indo-européennes, par le jeu combiné de trois facteurs : par le passage de l'accent au suffixe et la réduction de la racine devenue atone (état monosyllabique, *bhér tó* « ce qui est porté » ; état agglutinant, *bhertó ;* état de flexion, *bhrtó*) ; secondement, par l'infixation (mais cet élément est encore à peu près insaisissable) ; enfin, analogie généralisant certains types de flexion d'une façon irrégulière, comme par exemple, ont été créés *cêpî*, *fêcî*, sur le modèle du régulier *sêdî* (pour *se-sd-î* parfait redoublé). — Le fait est que le passage de l'agglutination à la flexion est encore fort difficile à déterminer ; mais, en somme, nous nous en tenons à cette définition précise que la flexion consiste en une modification phonique des racines, modification répondant à une variation de signification.

§ 2. Flexion indo-européenne et flexion sémitique.

Nous passerons en revue tout à l'heure, avec plus ou moins de détails, les deux systèmes de langue à flexion, le système indo-européen (sanscrit, perse, grec, latin,

(1) *Esquisses morphologiques*, I et II.

idiomes slaves, celtiques, etc.) et le système sémitique (hébreu, arabe, etc.). Mais, avant de procéder à cet examen, il nous faut mettre en relief un fait très important et d'ordre général.

Ce n'est point seulement par leurs racines que les langues sémitiques et les langues indo-européennes sont totalement distinctes les unes des autres ; elles diffèrent encore en ce qui concerne leur structure elle-même. Les unes et les autres sont bien des langues à flexion, elles ont dépassé la forme de l'agglutination pure et simple, mais il s'en faut de beaucoup que la flexion soit chez les unes ce qu'elle est chez les autres.

Schleicher (1) et Withney (2) ont examiné cette question de très près.

Avant que les idiomes sémitiques devinssent autant de langues bien distinctes, le système sémitique, dit Schleicher, ne possédait point de racines auxquelles on pût donner une forme sonore quelconque. Il se distinguait profondément en cela du système indo-européen. Le sens de la racine était attaché à de simples consonnes ; c'est en leur adjoignant des voyelles qu'on indiquait les relations du sens général. C'est ainsi que les trois consonnes *q t l* constituent la racine de l'hébreu *qâtal*, de l'arabe *qatala* « il a tué », *de qutila* « il fut tué », de l'hébreu *hiqtil* « il fit tuer », de l'arabe *maqtûlun* « tué ». Il en est tout différemment dans le système indo-européen, où le sens est attaché à une syllabe parfaitement prononçable.

Seconde différence. La racine sémitique peut admettre toutes les voyelles propres à modifier son sens. La racine

(1) *Die deutsche Sprache*, 2e éd., p. 21. Stuttgart, 1869. *Semitisch und Indogermanisch*, Beiträge zur vergleichenden Sprachforschung, t. II, p. 236. Berlin, 1861.

(2) *Language and the Study of Language*, 3e édit., 390. Londres, 1870.

indo-européenne, au contraire, possède une voyelle qui lui est propre, qui est organique, une voyelle fondamentale. Sans doute, dans le sanscrit *manvê* « je pense », dans le latin *mens, moneo*, dans le gothique *gamunan* se souvenir », nous trouvons comme voyelles du radical ici un *a*, là un *e*, ailleurs un *o* ou un *u* ; mais ces voyelles du radical sont loin d'être chacune la voyelle radicale vraie, la voyelle fondamentale. Les néo-grammairiens ont pu nier que cette voyelle fût un *a*, comme on l'a cru longtemps, mais, en somme, une voyelle fondamentale a existé. La voyelle organique de la racine indo-européenne ne peut d'ailleurs se changer, à l'occasion, qu'en telle ou telle autre voyelle, d'après des lois que reconnaît et détermine l'analyse linguistique.

Troisième différence. La racine sémitique est trilitère : *qtl* « tuer » *ktb* « écrire », *dbr* « parler », elle provient, nul n'en doute, de formes plus simples, mais enfin c'est ainsi qu'on la reconstitue. Par contre, la racine indo-européenne est bien plus libre de forme, comme le montrent, par exemple, *i* « aller », *su* « verser arroser » ; toutefois elle est monosyllabique.

Le système sémitique n'avait que trois cas et deux temps, le système indo-européen a huit cas et six temps au moins.

Tous les mots de l'indo-européen ont une seule et même forme : celle de la racine (modifiée ou non), accompagnée du suffixe dérivatif. Le sémitique emploie aussi cette forme (exemple, l'arabe *qatalta* « toi, homme, tu as tué »), mais il connaît aussi la forme où l'élément dérivatif est préfixé, celle où la racine est entre deux éléments dérivatifs, d'autres formes encore.

La flexion du système sémitique, dit W. Whitney, est totalement différente de la flexion indo-européenne et ne permet point de faire dériver les deux systèmes l'un de l'autre, non plus que d'un système commun. La caracté-

ristique fondamentale du sémitisme réside dans la forme trilitère de ses racines. Celles-ci sont composées de trois consonnes, auxquelles différents voyelles viennent s'adjoindre en tant que formatives, c'est-à-dire en tant qu'éléments indiquant les relations diverses de la racine. En arabe, par exemple, la racine *qtl* présente l'idée de « tuer » et *qatala* veut dire « il tua », *qutila* « il fut tué », *qatl* « meurtrier », *qitl* « ennemi », etc.

A côté de cette flexion due à l'emploi de différentes voyelles, le sémitisme forme aussi ses mots en se servant de suffixes et de préfixes, parfois également d'infixes. Mais l'agrégation d'affixes sur affixes, la formation de dérivatifs tirés de dérivatifs, lui est comme inconnue ; de là provient la presque uniformité des langues sémitiques.

La structure du verbe sémitique diffère profondément de celle du verbe indo-européen. A la seconde et à la troisième personne, il distingue le genre masculin ou féminin du sujet : *qatalat* « elle tua », *qatala* « il tua » ; c'est ce que ne font point les langues indo-européennes : sanskrit *bharati* « il porte, elle porte ».

L'antithèse du passé, du présent, du futur, qui est si essentielle, si fondamentale dans les langues indo-européennes, n'existe point pour le sémitisme : il n'a que deux temps, répondant, l'un à l'idée de l'action accomplie, l'autre à celle de l'action non accomplie.

On voit combien les différences de structure sont considérables entre les deux systèmes et combien leurs modes de flexion sont différents (1).

Ces deux familles de langues sont donc sorties par des voies toutes différentes de la phase agglutinative qu'elles ont dû traverser, et elles sont aussi indépendantes l'une de l'autre par leur structure, qu'elles le sont par leurs

(1) Th. Noeldeke, *Orient und Occident*, t. II, p. 375. Gœttingen, 1863.

racines, dont la prétendue réductibilité à d'anciennes formes communes ne mérite plus d'être débattue.

A. LES LANGUES SÉMITIQUES.

Nous allons parler à tour de rôle, sous trois rubriques, des langues sémitiques, des langues khamitiques, des langues indo-européennes.

Il est inutile, assurément, de faire remarquer combien les noms de *sémitisme* et de *langues sémitiques* sont conventionnels. Ils ont été inventés pour cadrer avec l'ethnographie du Testament hébraïque, et, en fait, ils ne s'accordent cependant pas avec le récit de la légende. Celle-ci regarde comme descendants du Sem biblique, des individus dont la langue ne saurait être classée parmi celles que nous appelons « sémitiques » et ne considère point comme descendants du même auteur des populations dont la langue est incontestablement sémitique. Quoi qu'il en soit, ces noms de sémitisme et de sémitique ont acquis aujourd'hui une telle notoriété qu'il n'y a plus à songer à leur substituer quelque autre dénomination plus acceptable. On emploie parfois le terme mieux justifié de *langues syro-arabes*, mais on ne peut penser à le faire prévaloir contre la dénomination courante et reçue. Comme le dit d'ailleurs E. Renan, dans son ouvrage aujourd'hui classique et auquel nous allons beaucoup emprunter (1), cette dernière dénomination ne peut avoir d'inconvénient du moment qu'on la prend comme une simple appellation conventionnelle et que l'on s'est expliqué sur ce qu'elle renferme de profondément inexact.

(1) *Histoire générale et système comparé des langues sémitiques*, première partie. *Histoire générale des langues sémitiques.*

§ 1. Du sémitisme en général et de l'ensemble des langues sémitiques.

Malgré les travaux de Gesenius (1786-1842) et malgré ceux de Ewald, il n'y a point encore de grammaire comparée des langues sémitiques, il n'existe pas d'ouvrage véritablement général sur l'ensemble du caractère de ces langues. On peut dire que c'est là une lacune considérable et qu'il serait fort important de combler.

Une fois ce travail mené à bonne fin, il faudrait l'entreprendre ensuite sur les langues dites khamitiques, puis comparer les formes khamitiques communes aux formes premières du sémitisme et esquisser les traits principaux de la grammaire khamito-sémitique. Cette grammaire tiendrait assurément en un bien petit nombre de pages, mais on ne peut guère douter de la possibilité d'une pareille entreprise.

Peut-être même pénétrera-t-on plus avant dans les secrets de l'évolution des langues à flexion et tentera-t-on de reconstituer les traits généraux qu'elles devaient offrir lorsqu'elles en étaient encore à la période de l'agglutination.

On a cherché déjà à ramener les racines sémitiques, qui, ainsi que nous l'avons dit, sont trilitères, c'est-à-dire composées de trois consonnes, à une forme bilitère. Il est permis d'assurer, sans témérité, que cette entreprise sera couronnée d'un heureux succès. On trouvera assurément un grand secours dans la comparaison avec les langues khamitiques, comme l'a dit Benfy (2) et comme

(1) Chavée, *les Langues et les Races*, p. 44. Paris, 1862. Renan, *Op. cit.*, liv. I, chap. III ; *Rapport annuel*, Journal asiatique, VIIe série, t. IV, p. 27. Paris, 1874. Schleicher, *Die Unterscheidung von Nomen und Verbum in der lautlichen Form*, p. 18.

(2) *Geschichte der Sprachwissenschaft und orientalischen Philologie in Deutschland*, p. 691. Munich, 1869.

il l'a démontré. Le soi-disant trilittérallisme primitif (Sayce, trad. franc., p. 66) est aujourd'hui insoutenable. Quant aux racines quadrilitères, elles se ramènent toutes aisément à des formes anciennes moins développées.

La tâche d'une grammaire comparée des langues sémitiques sera de déterminer l'emploi que l'on fit des différentes voyelles, dans le but de donner au nom ainsi constitué tel ou tel caractère. Ce mode de formation du nom est assez élémentaire. Il en existe un autre, celui de la dérivation. Ce dernier consiste en ceci, que certaines syllabes sont préfixées à la racine ou encore qu'elles lui sont suffixées. L'origine de ce dernier procédé de dérivation est moins ancienne dans le sémitisme que celle de la préfixation.

Dans la langue sémitique commune, le nom aurait connu les trois genres, masculin, féminin et neutre (1), mais ce dernier aurait disparu à une époque fort ancienne. Le masculin n'était exprimé par aucun élément spécial, à l'encontre du féminin qui était rendu, selon toute vraisemblance, par la terminaison *at* (2). Le *t* du féminin s'est maintenu dans l'ancien assyrien, dans le vieil arabe. Dans l'arabe moderne, si un *a* le précède, ce *t* se change en *h* quand il est à la fin du mot ; de même en hébreu. La caractéristique organique du pluriel était peut-être *mûn* (3), peut-être *umû*, *unû* (4) ; peut-être encore était-ce quelque autre forme. Cele du duel paraît en avoir procédé.

La déclinaison comportait trois cas, chiffre de beaucoup inférieur à celui des cas de la langue commune

(1) Ewald, *Ausführliches Lehrbuch der hebræischen Sprache*, 8e édit., p. 445. Gœttingen, 1870.
(2) *Ibid.*, p. 446.
(3) *Ibid.*, p. 465.
(4) Frédéric Müller, *Der Verbalausdruck in semitischen Sprachkreise*. Sitzungsberichte der phil. hist. Classe der k. Akademie der Wissensch., t. LX, p. 520. Vienne, 1868.

indo-européenne. Ces trois cas étaient le nominatif, le génitif, l'accusatif ; toute la famille sémitique, à l'exception de l'arabe, les a perdus en grande partie, ainsi que nous le verrons en traitant séparément de chacun de ces idiomes. D'après certains auteurs, la voyelle *u* aurait été le signe du nominatif, *i* celui du génitif (en principe), *a* celui de l'accusatif (Olshausen, *op. cit.*, p. 25 ; cf. Ewald, *op. cit.*, p. 523). D'après Fr. Müller, *u* du nominatif tirerait son origine du pronom *hu* « il » de la troisième personne ; le signe de l'accusatif aurait été *an*, démonstratif, devenu bientôt *a* ; le *i* du génitif aurait la même origine que le pronom relatif (*Grundriss*, tome III, 2e partie, p. 343 ; voir encore sur les cas, dans le sémitisme, *Journal asiat.*, tome X de la 6e série, p. 373).

Quant aux pronoms personnels, on n'est pas encore arrivé à restituer leurs formes communes ; ce serait là cependant un point très important à atteindre (1).

D'après Fr. Müller, il faudrait les restituer comme suit :

Je	*'an-â-ki.*	Nous	*'an-â-knû.*
Tu (masc.).	*'an-tâ.*	Vous (masc.).	*'an-t-umû.*
Tu (fém.)..	*'an-ti.*	Vous (fém.)..	*'an-t-in.*
Il	*huwâ, tuwâ.*	Ils	*h-umû, t-umû*
Elle	*hiyâ, tiyâ.*	Elles	*h-in, t-in.*

Le pronom possessif que l'on postpose au substantif (arabe *kitâb-i* « mon livre » ,hébreu *sûs-i* « mon cheval »), remonte au pronom personnel d'une façon évidente (sur les pronoms sémitiques, voir *Journal asiat.*, tome XII de la 6e série, p. 67).

La langue sémitiquée commune ne connaissait que deux temps : un temps *parfait* indiquant l'action accomplie, un temps *imparfait* indiquant l'action non accomplie.

(1) Chavée, *Op. cit.* Fr. Muller, *Op. cit.*, p. 353.

Les deux temps se distinguent l'un de l'autre par la position qu'occupe à côté du thème le suffixe personnel. Ce suffixe (par exemple TA de la seconde personne masculine du singulier) est-il placé après le thème, l'action est accomplie, le temps est parfait : KATABATA « tu as écrit », arabe KATABTA ; est-il placé avant : TAKATABA, arabe *taktuba*, le temps est imparfait, l'action n'est pas encore accomplie : temps aoristique, temps duratif.

Le verbe sémitique organique aurait pu, d'après Fr. Müller, se conjuger d'après treize formes thématiques : la forme simple, *qatala* « il tua », une forme intensive ou renforcée, *qattala* ,etc. Voici, d'ailleurs, en partie, le tableau de cette restitution :

1	forme	fondamentale	*qatala.*
2	—	intensive	*qattala.*
3	—	d'influence	*qâtala.*
4	—	causative de 1.....	*'a-qtala.*
5	—	causative de 2.....	*'a-qattala.*
6	—	causative de 3.....	*'a-qâtala.*
7	—	réflexive de 1.....	*ta-qatala.*
8	—	réflexive de 2.....	*ta-qattala.*
9	—	réflexive de 3.....	*ta-qâtâla.*
	Etc.,	etc.	*etc.*

Aucun des idiomes sémitiques n'a conservé ces treize formes ; l'arabe en possède neuf ; l'hébreu cinq (*qâtal. qittêl, hiqtil, hithqattêl, niqtal*) ; l'araméen en possède six.

Le passif serait une forme réflexive formée à l'aide de l'élément pronominal *u* (*hu*) :

1	forme	organique fondamentale.....	*qûtala.*
2	—	intensive	*quttala.*
3	—	d'influence	*qutala.*
—	—	causative de 1.....	*'u-qtala.*

Le verbe sémitique incorpore le pronom régime (voir

ce qui a été dit à ce sujet à l'occasion des langues ouralo-altaïques) : *qatala-ni* « il me tua » ; *qatalnâ-hu* « nous le tuâmes ».

Quant à la numération, elle est décimale.

Le système général de l'alphabet sémitique aurait été emprunté aux hiéroglyphes égyptiens (1). On a cru longtemps que cet emprunt était dû aux Phéniciens, mais cette opinion, paraît-il, est inexacte. D'après Ewald, il aurait été fait par une autre nation sémitique dont les rapports avec l'Egypte auraient été plus intimes encore. En tout cas on ignore le nom du peuple qui rendit à la civilisation cet immense service de convertir les anciens hiéroglyphes en écriture alphabétique. Les figures, les images de l'écriture égyptienne étaient converties en autant de signes représentant tel ou tel son et elles allaient perdre peu à peu leur ancien caractère de purs et simples dessins.

Le vieil alphabet sémitique se compose de vingt-deux consonnes. Chacune d'elles devait exprimer l'articulation correspondant à l'articulation initiale de l'être ou de l'objet représenté par le signe lui-même. Ainsi l'ancienne image du chameau représentait un *g* dans l'écriture alphabétique des Sémites, vu que le nom du chameau (chaldéen *gimel*, syriaque *gomal*) commençait chez eux par un *g*. Il est à peine utile d'ajouter que ces nouveaux signes ces signes alphabétiques se modifièrent diversement chez les différents peuples qui les adoptèrent.

On divise d'habitude l'écriture sémitique en trois groupes distincts.

Le groupe occidental comprend le système du phénicien et de l'ancien hébreu. Cette vieille écriture hébraïque était encore en usage au second siècle avant notre ère.

A l'est, dans les régions de l'Euphrate et du Tigre, les

(1) E. de Rougé, *Mémoire sur l'origine égyptienne de l'alphabet phénicien*. Paris, 1874.

formes de l'ancien alphabet sémitique s'étaient arrondies. On y vit bientôt naître une écriture cursive qui se répandit dans le pays de l'ouest et le nord de l'Arabie.

Dans le sud de cette dernière contrée, le système himyaro-éthiopien s'était développé.

Nous dirons quelques mots de ces différentes variétés de l'écriture sémitique en traitant des divers idiomes de cette même famille.

Quant à l'écriture cunéiforme assyrienne, avec laquelle sont rédigés les textes de la troisième colonne des inscriptions achéménides, nous devons lui assigner une tout autre provenance. C'est ce que nous verrons également en temps opportun.

La classification des langues sémitiques est aujourd'hui assez bien fixée. Elle était loin d'être aussi facile à établir que celle des langues indo-européennes. Les dialectes sémitiques, en effet, ne présentent point entre eux de ces différences caractéristiques qui séparent, par exemple, les langues celtiques des langues éraniennes, les langues italiques des langues slaves. L'on a pu dire, avec juste raison, que les différents idiomes sémitiques n'étaient pas plus éloignés les uns des autres que ne le sont les différentes langues d'une même branche dans la famille indo-européenne, par exemple le russe, le tchèque, le croate, ou encore l'anglais, le flamand, le danois.

On divise ordinairement les langues sémitiques en trois groupes distincts :

Le GROUPE ARAMÉO-ASSYRIEN, comprenant l'*assyrien* et les deux dialectes araméens, soit le *chaldéen* et le *syriaque ;*

Le GROUPE CHANANÉEN, comprenant l'*hébreu* et le *phénicien ;*

Le GROUPE ARABE, comprenant l'*arabe* proprement dit et les idiomes de l'Arabie méridionale, *himyarite*, et *chkili*, *ghez* et *tigré*, *amharique* et *harari*.

Quelques auteurs réduisent encore cette classification à deux groupes. Les deux premiers groupes n'en formeraient qu'un seul auquel ils donnent le nom de groupe septentrional, par opposition au groupe méridional formé des deux variétés du groupe arabe.

Nous allons jeter un coup d'œil sur ces différentes langues et nous chercherons, pour terminer, s'il n'est point possible de former quelque conjecture sur le lieu ou la forme commune dont elles procèdent toutes aurait été parlée.

§ 2. Groupe araméo-assyrien.

I. *Chaldéen et syriaque.*

On donne le nom de langue araméenne à deux dialectes de ce groupe fort rapprochés l'un de l'autre : le *chaldéen*, dialecte oriental ; le *syriaque*, dialecte occidental.

Le premier de ces dialectes s'étendait sur la plus grande part, sinon sur la totalité de la Babylonie et de l'Assyrie ; le second sur la Mésopotamie et la Syrie.

L'araméen se distingue spécialement par ce fait qu'il a fort mal conservé les anciennes voyelles sémitiques. On peut dire qu'il doit cette infériorité relative, par rapport aux autres idiomes sémitiques, à son développement plus précoce.

Les deux dialectes araméens ainsi que nous l'avons dit, sont fort peu différents l'un de l'autre. Leur mode d'accentuation est toutefois assez divergent : tandis qu'en principe, l'accent, en chaldéen, tombe sur la dernière syllabe et n'effecte l'avant-dernière qu'en certains cas déterminés, en syriaque, par contre, il tombe régulièrement sur l'avant-dernière syllabe et le cas où il doit

tomber sur la dernière ne forment que l'exception. Quant aux diversités purement grammaticales elles sont minimes ; le syriaque, par exemple, met souvent la voyelle *o* là où le chaldéen dit *a* ; ce dernier évite davantage les diphtongues. En somme cela est de peu d'importance.

L'araméen fut surtout un idiome populaire et les Juifs transportés à Babylone se familiarisèrent promptement avec lui. Ils le rapportèrent en Palestine.

L'ancien araméen ne nous a point laissé, comme l'a fait l'assyrien, dont nous parlerons tout à l'heure, de documents indigènes. C'est dans les livres saints des Juifs que nous trouverons les plus anciens textes araméens auxquels on donne le nom de « chaldéen biblique » et qui peuvent remonter au cinquième ou au sixième siècle avant notre ère. D'autres passages de la Bible, rédigés également en araméen, datent d'un âge moins reculé. Vers l'époque chrétienne apparaissent les Targums, traductions et paraphrases des livres juifs. La langue des Talmuds, plus vieille de quatre ou cinq siècles, est beaucoup plus chargée d'éléments étrangers à l'araméen et empruntés aux idiomes voisins.

Dans son histoire des langues sémitiques, E. Renan traite successivement de l'aramaïsme païen et de l'aramaïsme chrétien.

C'est dans le *nabaéten* et le *mendaïte* que se montre le premier. Le nom de langue nabatéenne équivaut à celui de chaldéen. Il ne nous reste de l'importante littérature de cet idiome que le traité de l'agriculture nabatéenne traduit en arabe au dixième siècle de notre ère, mais on ignore l'époque de sa rédaction originale. Le sabien, ou pour parler plus correctement, la langue des Mendaïtes, est loin d'avoir produit une littérature aussi considérable que paraît l'avoir été la littérature nabatéenne. Ce que nous en possédons semble postérieur à l'islamisme.

On connaît spécialement le « Livre d'Adam », amas d'imaginations ridicules. E. Renan signale comme particularité du mendaïte la confusion et l'élision fréquente des gutturales, le changement des douces en fortes et des fortes en douces, des contractions nombreuses.

L'aramaïsme chrétien a pour langue le *syriaque*. Cette langue ne nous offre point de traces d'une littérature plus ancienne que les premiers siècles de notre ère. Il paraît cependant hors de doute que le syriaque a possédé une littérature vraiment nationale. Les inscriptions palmyréennes datent des trois premiers siècles ; quant aux auteurs syriens les plus anciens, ils datent de la seconde partie du deuxième. On attribue à ce siècle la version « pechito » de la Bible, qui est le plus ancien ouvrage syriaque. Du quatrième au neuvième siècle fleurit une fort importante littérature de l'araméen chrétien, qui toutefois est singulièrement empreinte d'hellénisme. Elle servit en quelque sorte d'intermédiaire entre la science arabe et opéra la transition de l'une à l'autre. Presque toutes les traductions d'auteurs grecs en arabe auraient été faites, dit l'auteur que nous suivons, par des Syriens et sur des versions syriaques (1).

Au dixième siècle arrive la décadence, l'islamisme fait décidément prévaloir sa culture et le syriaque passe à la simple condition d'idiome liturgique. Il n'est plus guère parlé aujourd'hui que dans un très petit nombre de localités aux environs du lac d'Ourmia (un peu à l'ouest de Tabris, dans l'Aderbaïdjan), et l'on peut prévoir que dans un temps peu éloigné le néo-syriaque aura disparu (voir Spiégel, *Erân*, p. 43 ; Socin, *Sur les dialectes syriaques modernes*, Congrès des Orientalistes, 1876, tome II, p. 260 ; *Zeitschrift der deutschen morgenl. Gesellschaft*, tome XXI, p. 183).

(1) Renan, *Op. cit.*, liv. III, chap. III, § 2.

L'écriture syriaque arrondit considérablement les formes du vieux système d'écriture sémitique. Originairement elle ne figurait que les consonnes et laissait de côté toutes les voyelles ; ce n'est qu'à partir du septième et du huitième siècle que ces dernières y furent représentées définitivement.

II. *Assyrien.*

A côté du chaldéen et du syriaque se place l'*assyrien.* C'est la seconde langue du groupe nord-oriental des idiomes sémitiques, c'est la langue dans laquelle est rédigé le texte de la troisième colonne des inscriptions cunéiformes. Nous avons vu plus haut que l'assyrien aurait pu recevoir anciennement le nom « d'accadien » (nom qui a été donné par Hincks à l'idiome encore contesté que Jules Oppert appelle « sumérien » et regarde comme une langue touranienne).

L'assyrien n'a pas été reçu sans difficulté dans la famille des langues sémitiques. Il a fallu de longues et vives luttes pour le faire admettre à la place qu'il doit occuper légitimement et qu'on ne pourrait plus lui contester aujourd'hui. L'opposition qu'il a rencontrée a singulièrement profité aux études dont il a été l'objet, et l'on peut dire à cette heure que l'on sait de sa grammaire la plus grande partie de ce qu'il sera jamais possible d'en connaître. Les importants travaux de H. Rawlinson ont clos d'une façon définitive l'ère des écrits dont le but était de fixer la nature même de la langue assyrienne. Les objections durent tomber les unes après les autres, celle-ci la première, qui consistait à nier le caractère sémitique de l'assyrien, vu la diversité de son alphabet d'avec l'alphabet sémitique ordinaire.

Les différentes formes de l'écriture assyrienne (ninivite, babylonienne) sont composées d'espèces de coins

plus ou moins grands, disposés d'autre façon que ceux de l'écriture perse dont nous aurons à nous occuper en traitant des langues éraniennes. Ces caractères cunéiformes proviennent d'anciens caractères hiéroglyphiques qu'il n'est point difficile de reconnaître sous quelques-uns d'entre eux.

Les cunéiformes assyriens diffèrent des cunéiformes perses, mais ils sont à peu de chose près les mêmes que ceux employés dans la seconde colonne des inscriptions achéménides. Leur origine commune est évidente et se reconnaît du premier coup d'œil.

Ces caractères, dont le nombre est considérable, figurent ou des idées ou des sons. Les signes phonétiques, ceux qui représentent des sons, figurent des syllabes entières, et non pas telle ou telle consonne, telle ou telle voyelle Dès 1849, Hincks signalait ce fait que l'écriture assyrienne était syllabique.

On la transcrit facilement en caractères latins, mais, bien entendu, il ne peut en être de même des signes idéographiques. En effet, la valeur phonique de l'idéogramme ne peut être révélée que par des renseignements accessoires. Pour tourner la difficulté on est convenu de transcrire ces derniers caractères tout comme s'ils étaient des signes phonétiques, seulement on emploie dans cette transcription des lettres latines capitales.

Les textes assyriens que possèdent les musées d'Europe sont déjà fort nombreux et il est certain que leur quantité augmentera encore dans une très grande proportion. On rencontre dans le pays même une foule de monuments gravés parmi lesquels il en est de considérables. La troisième colonne des inscriptions des Achéménides est, comme l'on sait, rédigée en assyrien ; nous avons dit quelques mots sur la langue de la seconde colonne, p. 191, et nous parlerons à son temps de la langue perse, qui était celle de la première colonne.

J. Oppert, dont les travaux ont contribué en une mesure notable au déchiffrement des cunéiformes assyriens (1), peut être appelé à juste titre le fondateur de la grammaire assyrienne (2). Ses écrits ont marqué une période nouvelle dans l'assyriologie. D'autres grammaires ont paru depuis ses écrits, et l'étude de l'assyrien n'offre plus aujourd'hui de difficulté considérable (3).

Disons quelques mots ici de la grammaire assyrienne. Sa phonétique semble moins altérée que celle des deux dialectes araméens ; les sifflantes notamment sont soumises à moins de variations.

L'élément *at* (parfois *it*) est en assyrien, comme dans les autres langues sémitiques, le signe du genre féminin: *sar* « roi », *sarrat* « reine », *ilu* « dieu », *ilat* ou *ilit* « déesse » ; *rabu* « grand », *rabit* « grande ».

Le pluriel des noms masculins est en *i* (*în* en aranéen, *îm* en hébreu). Exemple : *yûm* « jour », *yûmi* « jours ». Dans les noms féminins le pluriel, en principe, est en *ât* (*ôt* en hébreu). Parfois il est en *ût*, parfois en *it*.

Quant au duel, il ne se présente que très rarement.

En assyrien, comme en arabe, on trouve les cas que les autres idiomes congénères ont perdus : *sarru* « rex », *sarra* « regem », *sarri* « regis » ; *sarratu* « regina », *sarrata* « reginam », *sarrati* « reginæ ». Au pluriel *sarrî* « reges » pour les trois cas; *sarrâtu* « réginæ », acc. *sarrâta*, gén. « *sarrâti* ». Les formes casuelles en *u*, *a*, *i* sont postérieures à des formes en *um*, *am*, *im* : *yûm* « dies », accus. *yumam*, génit. *yûmi*. Ce phénomène de

(1) *Expédition scientifique en Mésopotamie*, t. II, Paris, 1859.
(2) *Eléments de la grammaire assyrienne*, 2e édit. Paris, 1868.
(3) Ménant, *Exposé des éléments de la grammaire assyrienne*. Paris, 1868. — *Le Syllabaire assyrien*. Paris 1869-74. *Leçons d'épigraphie assyrienne*. Paris, 1873. — Sayce, *An Assyrian Grammar*. Londres, 1872. — Schrader, *Die assyrisch-babylonischen Keilinschriften*, Zeitschr. der deutschen morgenlændischen Gesellschaft, t. XXI, p. 1-392. Leipzig, 1872.

la « mimmation » serait l'équivalent du phénomène de la « nounnation » que nous constaterons plus bas dans la langue arabe. Par la suite des temps le *m* terminal des anciennes désinences assyriennes disparut peu à peu, et la voyelle elle-même ne fut plus régulièrement respectée. Nous constaterons ci-dessous que dans l'arabe littéral les voyelles *u*, *i*, *a* terminent le nom, s'il est précédé de l'article, et qu'au contraire, si le nom n'est pas précédé de l'article, la désinence est nasale. Quant à l'assyrien, il n'a point d'article.

De même que les autres langues sémitiques, il rend le pronom possessif en suffixant au nom un élément pronominal : *bitya* « ma maison », *babiya* « mes portes », *sumya* « mon nom », *sumiya* « mes noms ». Pour la seconde personne du singulier, le suffixe indiquant la possession est *ka* au masculin, *ki* au féminin ; de là : *sumka* « ton nom » (le nom de toi, homme), *sumiki* « tes noms » (les noms de toi, femme).

Nous avons dit que la langue sémitique commune possédait pour son verbe deux temps : un temps parfait indiquant l'action accomplie, un temps imparfait indiquant l'action non accomplie, et nous avons ajouté que, dans les formes du premier de ces temps, le pronom personnel était placé après le thème, tandis qu'il était placé avant ce même thème dans le temps imparfait. En assyrien, le système de formation est le même, mais il s'est opéré une évolution dans la signification même.

Quant aux pronoms régimes immédiatement accolés au verbe, nous les trouvons ici comme dans l'ensemble du système sémitique. Ainsi la phrase « je les ai soumis » se dit en un seul mot : à la forme qui signifie « j'ai soumis » ou suffixe le pronom *sunut* « eux ».

Ajoutons, pour terminer, que l'assyrien fut parlé presque jusqu'à notre ère. Depuis plusieurs siècles déjà l'araméen tendait à le supplanter. Il le supplanta, en

effet, mais disparut à son tour devant les progrès de la langue arabe.

§ 3. Groupe chananéen.

Les langues chananéennes sont mieux conservées, dans leur ensemble, que les idiomes araméens, ainsi que le montrent très clairement les formes de l'ancien hébreu, de l'hébreu classique.

I. *Hébreu.*

Il faut reconnaître dans l'hébreu, avec Ewald, trois périodes (1).

Les fragments qui datent de l'époque de Moïse nous font voir la langue hébraïque toute formée et essentiellement la même que celle des temps plus modernes. A cette époque, elle devait donc être déjà fort ancienne. Dans la seconde période, dès le temps des rois, elle tend à se différencier en deux sortes de style, l'un plus vulgaire, l'autre plus artistique. La troisième période commence au septième siècle avant notre ère ; c'est l'époque de la décadence, c'est l'époque à laquelle l'araméen s'étend de plus en plus.

Les différences, cependant, sont peu considérables entre chacune de ces périodes.

« Ce qu'il importe de maintenir, dit E. Renan, c'est l'unité grammaticale de la langue hébraïque, c'est ce fait qu'un même niveau a passé sur les monuments de provenances et d'âges si divers qui sont entrés dans les archives des Israélites. Sans doute, il serait téméraire

(1) *Ausführliches Lehrbuch der hebræischen Sprache*, 8e édit., p. 23. Gœttingen, 1870.

d'affirmer avec F. Movers qu'une seule main a retouché presque tous les écrits du canon hébreu pour les réduire à une langue uniforme. Il faut reconnaître, toutefois, que peu de littératures se présentent avec un caractère aussi impersonnel et ont moins gardé le cachet particulier d'un auteur et d'une époque déterminée. » *Op. cit.*, liv. II. chap. Ier.

C'est seulement à partir du onzième siècle avant notre ère que des écrits hébreux se présentent à nous sans avoir été remaniés postérieurement ; trois au quatre cents ans plus tard, la langue hébraïque entre dans son âge d'or, puis, vers le sixième siècle, elle commence à se perdre comme langue populaire.

Bien avant l'époque des Macchabées, l'araméen était devenu prépondérant en Palestine. On continue cependant à rédiger encore des livres en hébreu jusqu'à une centaine d'années environ avant notre ère.

E. Renan divise en deux périodes distinctes l'histoire de l'hébreu moderne, c'est-à-dire de l'hébreu post-biblique. La première s'étend jusqu'au douzième siècle et a pour monument principal la Michna, recueil de traditions rabbiniques, espèce de seconde Bible; on y rencontre un certain nombre de mots araméens hébraïsés, des mots grecs et des mots latins. Après avoir adopté, au dixième siècle, la culture arabe, les Juifs virent renaître leur littérature quand leurs compatriotes chassés de l'Espagne musulmane gagnèrent la France du Sud. La langue de cette époque est encore aujourd'hui l'idiome littéraire des Juifs.

Le système des voyelles hébraïques est des plus simples, comme celui de l'araméen. Le système des consonnes est riche en sifflantes et en aspirations, comme c'est le cas dans toutes les langues sémitiques. Les sifflantes sont au nombre de quatre : elles répondent à notre *ch* (de « chercher »), à notre *s* (« de sensé »), à notre *z*

et à une sorte de *s* assez proche du « ts » français. L'hébreu a donné à ses sifflantes, en général, une bien plus grande importance que ne leur en ont attribué les autres langues sémitiques. On compte également quatre aspirées ; deux d'entre elles sont assez douces; les deux autres, le « heth » et le « ghaïn », sont gutturales et permutent parfois avec *k*, *q*. Outre les trois paires d'explosives *k-g*, *t-d*, *p-b*, l'hébreu possède un *q* plus énergique que le simple « k », et un *th* (ainsi transcrit par certains auteurs), plus énergique que « t ». Il existe également une explosive labiale distincte de « p » et que l'on transcrit souvent *f*. Il est bon de noter, d'ailleurs, que les consonnes, qui par nature sont susceptibles d'être aspirées, le sont en réalité dans la prononciation lorsqu'elles se trouvent précédées d'une voyelle. L'hébreu possède, en outre, les vibrantes *r*, *l*, les nasales *n*, *m*, la demi-voyelle *y* (souvent transcrit « j » à la façon allemande) et *v*.

En principe, le féminin des noms est en *at*. Mais parfois le *t* final se change en une simple aspiration ; parfois il n'est pas précédé de *a*. C'est un sujet particulier qu'il faut étudier en détail dans les grammaires bien faites.

Comme signe du pluriel, les noms masculins s'adjoignent l'élément *im* (que remplace parfois la forme araméenne *în*), et les noms féminins l'élément *ôt*. Ici encore nous ne donnons que la règle très générale. Le suffixe du duel est *ayim* : *yâdhayim* « les deux mains ».

Nous avons parlé ci-dessus des suppositions qui ont été faites par différents auteurs pour restituer la forme primitive des cas sémitiques. Quoi qu'il en soit, il ne reste plus en hébreu que des traces fort douteuses de l'ancien suffixe du nominatif et il en est de même pour l'accusatif et le génitif.

Dans la pratique, le nominatif, ayant ainsi perdu la

désinence qui le caractérisait jadis, est rendu par la forme la plus simple du nom, par le thème lui-même.

Les autres cas sont exprimés par l'emploi de prépositions ou au moyen du procédé que l'on appelle l'*état construit.*

Affectant la forme de l'état construit (opposé à l'état absolu), un mot se trouve placé, par là même, vis-à-vis d'un autre mot, dans une véritable condition de dépendance.

On voit déjà que la fonction principale de l'état construit est d'exprimer l'idée du génitif. Au singulier, les noms masculins à l'état construit restent, en principe, tels quels et précèdent immédiatement le mot qu'ils gouvernent. Les pronoms possessifs (mon, ton, son, etc.) sont exprimés constamment par ce procédé syntaxique. Pour dire « son peuple », par exemple, on accolera le mot « lui » après le mot « peuple »; de là : « peuple-lui », c'est-à-dire le peuple de lui, son peuple, *gham ô ; ben î* « mon fils ». Au pluriel, à l'état construit, les noms masculins perdent la consonne terminale *m*, parfois même la voyelle qui précède.

Nous avons vu plus haut que le *t* final des noms féminins se changeait parfois en une aspiration ; à l'état construit, le *t* organique de ces mots féminins apparaît dans toute sa rigueur. Au pluriel, les féminins gardent leur désinence *ôt.* Nous n'indiquerons ici, bien entendu, que les règles générales de l'état construit; dans les grammaires spéciales il comporte des explications assez minutieuses qui ne peuvent nous arrêter en ce moment.

En employant comme il le fait des prépositions dont la fonction est de suppléer aux terminaisons qui indiquaient les cas, l'hébreu présente une physionomie tout à fait analytique. Il est inexact en somme de parler, avec les grammaires ordinaires, d'un datif, d'un locatif, d'un ablatif hébreux ; les formes auxquelles on se plaît

à donner ce nom ne sont autre chose que des composés d'une préposition et d'un nom ou d'un pronom. Certaines de ces prépositions, les plus usitées, ne sont formées que d'une simple consonne : *l* « à, vers », *b* « dans ».

On connaît l'origine de presque toutes ces particules, qui, à l'inverse des prépositions indo-européennes issues pour la plupart de pronoms, proviennent, en principe, de racines verbales.

La flexion joue un rôle important dans la formation des noms ; nous avons dit précédemment qu'elle consistait dans la variation, dans la variabilité des voyelles du mot. C'est affaire aux grammaires spéciales que d'énumérer ces changements ; nous n'avons ici qu'à renvoyer à ce que nous avons dit ci-dessus de la flexion en général.

L'hébreu possède un article qui se joint au nom d'une façon tout à fait intime et qui a pour fonction exclusive celle d'un simple déterminatif : c'est *hal-*, qui répond à l'arabe *'al-*. La consonne *l* s'assimile toujours à la consonne initiale du nom suivant et parfois la voyelle *a* s'allonge. Ainsi de *mâqôm* « lieu » on fait *hamma qôm* « le lieu » ; de *kôhên* « prêtre » on fait *hakkôhèn* « le prêtre ». Après certaines prépositions, l'aspiration *h* vient à tomber.

Nous avons dit précédemment que le système sémitique ne possédait que deux temps : un temps désignant l'action accomplie, un temps désignant l'action non accomplie. L'hébreu demeure fidèle à cette conception si simple. Ces deux temps, comme nous l'avons vu, se distinguent par la position du suffixe personnel. Dans le temps parfait il est placé après le thème ; il est placé avant dans le temps imparfait.

Ainsi dans *zâqantî* « je suis vieux, je suis vieille, j'ai vieilli », dans *hâlaktî* « je suis allé », *yâladtî* « j'ai enfanté », nous reconnaissons des formes du temps par-

fait, vu que l'élément pronominal (*tî*) est à la fin du mot. Au contraire, dans *nâsub* « nous retournerons », l'élément personnel est placé avant le thème et nous avons affaire au temps qui indique que l'action n'est pas accomplie.

Quant aux formes mêmes du verbe hébraïque, elles ne sont qu'au nombre de cinq. Nous avons dit plus haut que l'on pouvait en compter treize dans le sémitique commun. Ce sont, en hébreu, la forme simple, *qâtal*, la forme intensive *qittêl*, causative du simple, *hiqtîl*, réflexive de l'intensif, *hithqattêl*, réflexive du simple, *niqtal*. L'aranéen possède une forme dérivée de plus.

Jusqu'aux derniers siècles de l'ère ancienne, l'alphabet phénicien, raide et anguleux, était également l'alphabet hébraïque. Il fut remplacé, non sans avantage pour la commodité et la rapidité de l'écriture, par l'alphabet chaldéen, plus arrondi, plus suivi dans ses formes. L'ancien alphabet se retrouve sur des monnaies de l'époque des Macchabées et sur quelques pièces qui paraissent avoir été frappées plus tard encore, lors de la guerre contre les Romains. Cependant, au temps des Macchabées, il existait déjà chez les Hébreux une sorte d'alphabet plus récent qui demeura en usage chez les Samaritains (1).

L'alphabet nouveau, l'alphabet chaldéen, ne distinguait pas plus les voyelles que ne le faisait le vieil alphabet dit « phénicien ». C'était là une lacune considérable. On essaya bien de tirer parti de certaines consonnes pour figurer le son des voyelles, mais ce système, bien qu'appliqué avec une certaine critique, ne pouvait donner que des résultats très incomplets et peu satisfaisants. C'est aux Massorètes, dit-on, que l'on doit l'invention des points-voyelles ; elle daterait du commen-

(1) Olshausen, *Op. cit.*, p. 52.

cement du sixième siècle de notre ère. Un certain nombre de modifications utiles furent également apportées dans les signes des consonnes. L'on distingua, par exemple, les consonnes prononcées avec force d'avec les autres en plaçant un point au milieu de la lettre. Les sons « s » et « ch » étaient figurés par un seul et même caractère : un point diacritique placé sur ce signe, soit à droite, soit à gauche, lui donna tantôt la valeur de « ch » *s*, tantôt celle de « s ».

II. *Phénicien* (1).

On ne sait que fort peu de chose des populations qui occupèrent la Palestine avant les tribus sémitiques venues de l'Orient, peut-être du sud-est, qui s'appelèrent elles-mêmes chananéennes. Ces dernières, au nombre desquelles il faut compter les Phéniciens, durent céder devant la horde des Beni-Israël, qui sous la conduite de Josué, treize cents ans environ avant notre ère, envahit la plus grande partie de la Palestine.

Les Chananéens furent refoulés par cette invasion vers les régions maritimes de l'Ouest, et il est permis de croire que cet événement contribua d'une façon notable à développer leurs relations avec les côtes baignées par la Méditerranée. Les Israélites auraient rendu cette fois à la civilisation un service capital, bien qu'indirect.

Nous n'avons pas à nous occuper ici de la question de savoir si les Israélites émigrant d'un pays qui avait l'araméen pour langue, parlaient primitivement un dialecte araméen et s'ils ont plus tard emprunté leur langue aux Chananéens. Le fait seul qui doit nous frapper est celui de la presque identité du phénicien et de l'hé-

(1) Schroeder, *Die phœnizische Sprache*. Halle, 1869. L'un des meilleurs écrits sur le phénicien. Nous lui avons emprunté un certain nombre de renseignements. — Renan, *Op. cit.*, liv. II, chap. II.

breu. On peut dire sans crainte qu'il exista une langue chananéenne commune qui donna naissance, par la suite des temps, à l'hébreu et au phénicien. Ces deux idiomes sont frères, il faut les placer sur le même rang et l'on exprime une opinion tout à fait inexacte en disant, comme on le fait souvent, que le phénicien est un dialecte hébraïque.

Cette erreur remonte à l'époque où l'on chercha à interpréter pour la première fois les documents phéniciens. La grammaire comparée n'était pas encore connue à ce moment, et les philologues entre les mains desquels étaient tombés des textes phéniciens faisaient naturellement dériver cette langue de l'hébreu, avec lequel ils lui trouvaient une si frappante ressemblance. Le doute aujourd'hui n'est plus permis, et, ainsi que nous le disions tout à l'heure, les deux idiomes sont frères et descendent d'une mère commune. Une fois séparés l'un de l'autre, ils suivirent chacun leur propre destinée « et se développant à part chez des peuples opposés de caractères et de mœurs, ils devinrent avec le temps, comme le dit E. Renan, différents l'un de l'autre, non pour la grammaire, mais pour la physionomie générale du discours. » On a pu dire avec juste raison que leurs différences n'étaient que des provincialismes.

On cite, parmi les principales différences de l'hébreu et du phénicien, la propriété qu'avait ce dernier idiome d'employer dans le langage courant et usuel un certain nombre d'expressions et de formes qui passent en hébreu pour de purs archaïsmes ou ne sont usitées que dans le style élevé. Nombre de mots phéniciens ont une acception diverse de leurs correspondants en hébreu ; tantôt c'est un sens plus large, tantôt c'est un sens plus restreint. Le phénicien possède, d'autre part, une forme de pronom relatif plus primitive que la forme hébraïque et se distingue encore par quelques autres particularités,

assez bien connues aujourd'hui, mais que nous n'avons pas à énumérer ici dans leurs détails.

Le phénicien, tel que nous le connaissons par ses inscriptions, qui ne sont pas d'une très haute antiquité, présente des marques importantes d'aramaïsme, davantage, peut-être, que ne le fait l'hébreu. Le phénicien des colonies établies sur la côte nord de l'Afrique offre également ces traces d'aramaïsme, mais ce fait n'a rien de surprenant si l'on songe à la haute antiquité de l'influence araméenne et aux rapports constants qu'entretenaient les colonies avec la mère patrie.

Le *punique*, ou phénicien d'Afrique, notamment la langue des Carthaginois, se divise d'une façon assez tranchée en deux dialectes, l'un plus ancien, l'autre plus récent. L'ancien punique est identique au phénicien de Palestine. Le néo-punique est plus altéré et son orthographe est souvent vicieuse. Les monuments qui en sont restés proviennent particulièrement de la Tunisie et de l'Algérie orientale (1). L'alphabet néo-punique diffère notablement de l'ancien alphabet phénicien, dont il n'est d'ailleurs qu'une altération. Les caractères y sont, en général, assez simplifiés, et il en est qui se trouvent réduits à une simple ligne et se confondent presque les uns avec les autres.

On ne connaît la littérature phénicienne que par quelques fragments de l'Histoire phénicienne de Sanchoniaton et le Périple d'Hannon traduit en grecs; par des mots cités dans les auteurs anciens; par un passage de Plaute, puis par une série de monnaies et d'incriptions. Ces derniers monuments ont été découverts sur un grand nombre de points du littoral de la Méditerranée : à Marseille, en Espagne, sur la côte nord-africaine, dans les

(1) Judas, *Etude démonstrative de la langue phénicienne et de la langue libyque*. Paris, 1847. Du même auteur : *Nouvelles études sur une série d'inscriptions numidico-puniques*. Paris, 1857.

îles de Chypre, de Sardaigne, de Malte. Quant à la Phénicie, elle n'a fourni jusqu'à présent qu'un nombre assez restreint d'inscriptions.

Le phénicien disparut de la Palestine avant que le punique eût été absorbé, lui aussi, par des idiomes plus heureux. On peut penser avec E. Renan que le punique fut parlé jusqu'à l'invasion musulmane, et que la facilité avec laquelle l'arabe prit possession de certaines contrées de l'Afrique septentrionale, tint précisément à la persistance de l'idiome sémitique phénicien, dont l'arabe lui-même n'était pas fort éloigné, bien qu'il appartînt à une autre branche des langues sémitiques.

§ 4. Groupe arabe.

C'est à défaut d'autre nom que l'on donne celui d'arabe à la branche méridionale des idiomes sémitiques. Le mot d' « arabe » ne s'applique, à proprement parler, qu'à l'un des deux rameaux de cette branche, le rameau israélite. L'himyarite, le ghez et les autres idiomes sémitiques de l'Arabie méridionale, n'ont été bien connus que longtemps après l'arabe, et c'est en se fondant sur leur parenté très proche avec cette dernière langue, qu'on leur a appliqué, d'une façon un peu abusive, le terme générique d'arabe.

I. *Arabe.*

L'étonnante fixité propre aux idiomes sémitiques n'est nulle part plus manifeste que dans la langue arabe. Rien de plus curieux, on pourrait dire rien de plus étrange, que la constance presque parfaite de l'arabe à travers les temps qu'il a parcourus et dans les espaces immenses qu'il a occupés.

Dès l'époque de Mahomet (fin du sixième siècle et

commencement du septième) et même dans les poèmes antérieurs à l'islamisme, l'arabe apparaît tel qu'il est aujourd'hui encore dans l'usage littéraire, en possession de toutes ses formes, de son riche vocabulaire, et, l'on pourrait dire, dans sa perfection.

La forme primordiale du Coran était différente de celle des autres livres religieux. Selon l'expression de E. Renan, le Coran est comme le recueil des ordres du jour de Mahomet. Le livre ne fut pas écrit tout entier au temps même du Prophète ; certains fragments sont un peu postérieurs. Quoi qu'il en soit, ses disciples compilèrent tous les fragments de l'enseignement du Prophète et en rédigèrent une espèce d'exemplaire typique, dont les copies furent révisées à leur tour, au milieu du septième siècle, sous le kalife Othman (644-656). La prépondérance du dialecte koreichite, parlé au centre même de l'Arabie, était définitivement établie. Quant au style du Coran, on sait qu'il est de deux sortes : la première partie est une sorte de prose poétique, la seconde partie est rhythmée.

Les poèmes antérieurs à l'islamisme n'ont pas dû le précéder de beaucoup. La langue des « moallakâts », que l'on ne fait remonter qu'au commencement du sixième siècle, est du pur arabe littéraire, ce n'est pas une forme plus ancienne de l'arabe.

Avant le commencement du sixième siècle, les Sémites de l'Arabie centrale ne connaissaient pas l'écriture proprement dite. Le système d'écriture arabe provient, comme l'on sait, de l'araméen. Dès ses premiers temps, l'alphabet arabe était fort imparfait ; un certain nombre de ses consonnes se trouvaient représentées par un seul et même signe, ce qui prêtait à la confusion. Il fut réformé d'assez bonne heure; dès le premier siècle de l'hégire, pense-t-on. D'ailleurs, cette réforme n'eut pas lieu tout d'un coup. Elle fut graduelle et amena peu

à peu l'alphabet arabe à la forme que nous lui connaissons actuellement, pourvu, comme il l'est, de signes accessoires indiquant les voyelles et d'autres signes ayant pour rôle de distinguer l'un de l'autre des caractères dont la forme était primitivement la même.

La transcription des lettres arabes en caractères latins a donné lieu à plusieurs systèmes. Au point de vue de l'étude linguistique, cette transcription est fort utile, et le procédé de R. Lepsius n'ayant pu s'imposer universellement, il serait bon d'étudier à nouveau cette question. Un point important est de ne jamais transcrire par deux lettres latines un seul et même caractère, comme le font la plupart des grammaires. (Voir *Journal asiat.*, 6e série, t. II, p. 60, 136.)

L'arabe possède une importance capitale dans l'étude comparée des langues sémitiques.

Nous avons déjà dit que l'arabe avait conservé les trois cas du sémitisme commun, nominatif, accusatif, génitif, dont on ne pouvait plus découvrir que de faibles traces dans les idiomes du Nord. Les copulatives de ces trois cas sont, ainsi que nous l'avons vu également, les trois voyelles *u*, *a*, *i*. Ces trois voyelles terminent le nom, si celui-ci est précédé de l'article (*alᶜ abdu* « servus », *alᶜ abda* « servum », *alᶜ abdi* « servi »); elles sont au contraire, suivies d'une nasale, si l'article n'est pas joint au nom, en autres termes, si le nom est indéterminé. — L'état construit, dont nous avons parlé à propos de l'hébreu, se retrouve aussi en arabe.

Le pluriel peut être rendu de deux manières. L'une d'elles appartient au système sémitique général. C'est l'adjonction au nom d'un nouvel élément: aux noms masculins *ûna* pour le nominatif, *îna* pour les autres cas (*ᶜabdûna* « serviteurs », gémit, *abdîna ;* avec l'article *ᶜal-abdûna*, *ᶜal-ᶜabdîna*) ; aux noms féminins *âton* pour le nominatif, *atin* pour les autres cas. (Nous avons dit

qu'en araméen le signe du pluriel était *în* pour les masculins, *ât* pour les féminins ; en hébreu *îm* et *ôt*.) On donne à cette forme de pluriel les noms de pluriel sain, pluriel parfait, pluriel externe, pluriel régulier.

L'autre forme reçoit les noms de pluriel brisé, rompu, imparfait, interne, irrégulier. Ici le pluriel est exprimé par une modification du thème : « Frangitur forma singularis vel mutata una alterave vocalium, vel aliqua literarum transposita aut abjecta, vel nova litera incerta (1). » Parfois on a recours à un allongement dans l'intérieur du mot que l'on fait précéder en même temps d'un *a* : *tifl* « enfant », *atfâl* « enfants » ; parfois encore on a recours à d'autres procédés. On en trouvera l'énumération dans les grammaires spéciales.

On forme le duel par l'addition de l'élément *âni* pour le nominatif, *aini* pour le cas oblique : *yadâni* « les deux mains ».

L'arabe possède les deux temps du système sémitique commun : le temps parfait indiquant que l'action est accomplie, et le temps imparfait indiquant que l'action n'est pas accomplie. Notre présent est rendu tantôt par le parfait arabe, tantôt par l'imparfait. On use, par exemple, du parfait si l'action présente a déjà été accomplie auparavant et si elle est une action continue, comme dans cette formule : « dixerunt dicuntque ». Au contraire, on emploie l'autre temps si l'action présente se lie avec une action dont il va être parlé immédiatement après. Il en est de même pour l'expression de notre futur. L'arabe le rend par le temps parfait s'il considère l'action à venir comme étant d'ores et déjà un fait acquis ou s'il forme le vœu qu'elle se réalise ; dans les autres hypothèses, au contraire, il emploie le second temps.

(1) Zschokke, *Institutiones fundamentales linguæ arabicæ*, Vienne, 1869. — H. Derenbourg, *Essai sur les formes de pluriels en arabe* (*Journal asiatique*, 1867). — Stan. Guyard, *Nouvel essai sur la formation du pluriel brisé en arabe*.

Quant à la formation même des deux temps, elle a lieu, comme dans les autres langues sémitiques, par ce fait que l'élément personnel, la syllabe indiquant la personne, prend place soit devant le thème, soit après.

Des treize formes thématiques du sémitisme général l'arabe, plus fidèle que l'hébreu, en a conservé neuf : *qatala*, *qattala*, *qâtala*, etc.

Ce serait une erreur que de regarder l'arabe vulgaire comme autre chose que de l'arabe littéraire simplifié. La distinction capitale entre les deux formes, la forme littéraire et la forme courante, c'est que la seconde a laissé tomber les cas qui sont conservés dans la première.

Les cas se reconnaissent dans l'arabe vulgaire par la position respective des mots ou par l'emploi de prépositions. Il en est donc arrivé au degré d'analytisme qui distingue également le syro-chaldaïque, l'hébreu, le phénicien. Au pluriel, la terminaison générale est *in* pour les masculins, *ât* pour les féminins.

L'article *'al* assimile fréquemment sa consonne ; dans la prononciation, il perd souvent sa voyelle.

Quant aux désinences des formes verbales, elles ont également souffert : l'arabe littéral dit *qatalta* « tu as tué », *qatallum* « vous avez tué » ; l'arabe vulgaire dit *qatalt*, *qatallu*, *teqtol* « tu tueras », *teqtolu* « vous tuerez ».

En tout cas, ainsi que le fait remarquer E. Renan (*op. cit.*, liv. IV, chap. II), nombre de faits démontrent que les procédés caractéristiques de la langue littéraire étaient usités dans l'ancienne langue arabe. C'est ainsi, par exemple, que les flexions propres à l'arabe littéral sont absolument nécessaires pour expliquer la métrique des vieilles poésies. On prétend même que certaines tribus de l'Arabie centrale observent encore aujourd'hui, dans le langage courant, les flexions qui n'appartiennent plus qu'à la langue écrite (*ibid.*). Pourtant, l'on serait

stement taxé de prétention et de pédanterie en se servant à bon escient de ces désinences finales dans le langage courant.

Il ne saurait être question de dialectes dans l'arabe littéral. C'est une langue fixée et qui devra s'éteindre sans rejetons. L'on ne peut en dire autant de l'arabe vulgaire, de l'arabe parlé. Si peu qu'il diffère de la langue écrite, nous voyons qu'il en diffère précisément par un de ces changements qui constituent la vie même de bien des langues, à savoir le passage d'un état synthétique à un état analytique. L'arabe vulgaire vit lentement, mais il vit. De là ses dialectes différents.

On en compte quatre principaux : celui de Barbarie, ceux d'Arabie, de Syrie, d'Egypte. On s'accorde à regarder les trois derniers comme fort peu distincts l'un de l'autre ; ils ont chacun une certaine quantité de locutions propres, de termes particuliers, mais là s'arrête leur diversité. Le dialecte de Barbarie offre quelques divergences grammaticales ; elles ne sont pas assez considérables, cependant, pour que ce dialecte ne soit compris aisément dans toute l'étendue du territoire qu'occupent les autres dialectes. (Cf. *Journal asiat.*, 5e série, t. VI, p. 549 ; t. XVIII, p. 357.)

Le *maltais* a une origine arabe, mais n'est plus qu'un jargon plein de véritables barbarismes et que les mots d'origine étrangère ont fortement pénétré. Il en était de même du *mosarabe* du sud de l'Espagne, qui ne s'est éteint, paraît-il, qu'au siècle dernier.

L'arabe a fourni à certaines langues de l'Europe et de l'Asie un grand nombre de mots. Les langues éraniennes actuelles, le persan entre autres, ont admis dans leur vocabulaire une foule de mots arabes ; le turc lui en a emprunté un très grand nombre. Quelques-unes des langues de l'Inde moderne possèdent une quantité de mots de la même origine. Enfin, parmi les idiomes européens,

les langues novo-latines notamment l'espagnol et le portugais) lui ont fait nombre d'emprunts. Parfois ces emprunts sont directs, parfois ils sont indirects. En français, nous pouvons citer les mots « coton, tasse, zéro, chiffre, jarre, algèbre, cramoisi. »

II. *Langues de l'Arabie méridionale et de l'Abyssinie.*

La seconde branche du groupe arabe, appelée parfois branche « ioktanide », est composée de deux familles d'idiomes que l'on a été un certain temps avant de classer, non seulement dans le groupe arabe, mais encore parmi les langues sémitiques. Ce rameau méridional du système sémitique occupe en Asie le sud de l'Arabie, et en Afrique l'Abyssinie.

La vieille langue de l'Arabie du Sud était l'*himyarite*, que l'on connaît aujourd'hui par un bon nombre d'inscriptions. Cet idiome possède, comme l'arabe, la forme particulière des pluriels brisés, dont nous avons parlé un peu plus haut .

L'alphabet himyarite a donné lieu à des recherches fort intéressantes. Il est acquis aujourd'hui que cet alphabet dérive de l'ancienne écriture sémitique (qui a donné naissance, ainsi que nous l'avons vu, à l'écriture chaldéenne, à l'écriture arabe, en un mot à tous les alphabets sémitiques, sauf les cunéiformes assyriens).

La conquête islamite renversa la civilisation himyarite, et l'arabe s'étendit peu à peu, dans le sud de la péninsule, jusqu'au littoral de la mer des Indes et du golfe d'Aden. Pourtant la langue himyarite ne périt point sans laisser de traces. Dans l'extrême sud de l'Arabie, notamment dans la région du Mahrah, on a constaté, il y a une quarantaine d'années, l'existence de l'idiome *ehkili*, le-

quel, s'il n'est point un descendant direct de l'ancien himyarite, en est au moins fort rapproché.

Dès une époque très reculée les Sémites de l'Arabie méridionale avaient connu et colonisé la côte sud-ouest de la mer Rouge. Ce fut plusieurs siècles avant notre ère, mais à une époque qu'on ne saurait déterminer avec quelque exactitude. Ils y portèrent, avec leur civilisation, l'idiome que l'on connaît sous le nom de *ghez* — parfois aussi sous la dénomination fautive d' « éthiopien » — et dont les formes sont intimement liées à celles de l'himyarite. Le ghez est aujourd'hui une langue savante ; il n'existe plus comme idiome populaire, comme idiome courant. C'est particulièrement une langue liturgique.

On sait que le christianisme prit possession de l'Ethiopie vers le quatrième siècle. De ce même siècle date, selon toute vraisemblance, la traduction de la Bible en ghez ; nombre de versions d'autres livres juifs et chrétiens enrichirent la littérature éthiopienne. Elle possède un certain nombre d'ouvrages, traduits, pour la plupart, soit du grec, soit de l'arabe.

L'arrivée des Jésuites en Abyssinie fut le signal de la décadence. Ces redoutables apôtres, dont les Abyssins ne se débarrassèrent que trop tard, « attirant à eux toute l'instruction et hostiles à l'enseignement indigène, laissèrent le pays, quand ils le quittèrent, dans une profonde barbarie, dont il n'est pas sorti jusqu'à nos jours. » (Renan, *op. cit.*, liv. IV, chap. I).

Le ghez était une langue fort développée : il possédait, comme l'arabe, les pluriels brisés et conservait encore certaines désinences terminales qu'ont perdues l'hébreu et l'araméen. Il est avéré aujourd'hui que l'alphabet ghez a la même source que l'alphabet himyarite (sur ce dernier, on peut consulter *Journal asiat.*, 6e série, tome XX, p. 518).

A côté du ghez, qui n'est plus aujourd'hui, ainsi que

nous l'avons dit, qu'un idiome savant et un idiome liturgique, un certain nombre de langues sémitiques dépendantes du même rameau, et qui, pour ne point procéder directement du ghez, lui sont du moins alliées de fort près, sont encore parlées actuellement en Abyssinie. On en cite trois principales : l'*amharique* dans l'Abyssinie du sud-ouest ; le *tigré* au nord ; le *harari* au sud-est, par le 40e degré de longitude et le 10e de latitude. Ces idiomes sont peut-être greffés sur d'autres langues plus anciennes et appartenant à un autre système, mais leur grammaire est incontestablement sémitique et l'on ne pourrait en aucune façon les séparer du ghez.

§ 5. Individualité des langues sémitiques. Leur patrie primitive.

On s'est beaucoup plus occupé jusqu'à ce jour de découvrir un lien commun entre les langues indo-européennes et les langues sémitiques que de comparer ces dernières entre elles.

Il serait cependant fort important de rétablir, au moins dans ses traits généraux, la grammaire de l'idiome commun dont sont sorties toutes les langues sémitiques.

Ces différentes langues, ainsi qu'on a pu le voir par ce qui précède, sont peu divergentes les unes des autres. Il est permis de supposer que, dans l'état actuel des connaissances acquises, la tâche de restituer leur grammaire commune ne sera pas trop ardue. Elle le sera bien moins, en tout cas, que ne l'a été l'entreprise analogue tentée sur les idiomes indo-européens et qui a eu un si heureux succès. La dernière et heureuse tentative est celle de Fr. Müller (*Grundriss der Sprachwissenschaft*, tome III, 2e partie, p. 314).

Il est à peine besoin de faire remarquer que les auteurs qui ont le plus cherché à rattacher les langues sémitiques aux langues indo-européennes, n'ont jamais pensé à cette objection, pourtant si naturelle, qu'il leur fallait comparer non point l'hébreu ou l'arabe au zend, au sanskrit, au grec, mais bien la langue commune sémitique à la langue commune indo-européenne.

On peut dire que tous les rapprochements qu'ils ont cherché à établir, reposent toujours sur les étymologies, jamais sur la grammaire. Voilà qui les condamne d'une façon irrémissible.

L'étymologie pure et simple, nous l'avons assez répété, n'est pas une science. Il est facile avec l'étymologie de faire dériver l'un de l'autre les idiomes les plus dissemblables, non seulement le basque de l'irlandais, l'étrusque du tibétain, mais, ce qui n'est pas plus sérieux, l'hébreu du sanskrit, ou, à volonté, le sanskrit de l'hébreu.

La grammaire, comme l'a fort bien dit E. Renan, est ce qui constitue l'individualité d'une langue ; or, « il faut renoncer à chercher un lien entre le système grammatical des langues sémitiques et celui des langues indo-européennes. Ce sont deux créations distinctes et absolument séparées. » (*Op. cit.*, liv. V, chap II.). Nous avons déjà parlé ci-dessus, en traitant de la flexion en général (p. 203), de la différence profonde, radicale, qui existe entre la grammaire sémitique et la grammaire indo-européenne ; nous n'y voulons pas revenir.

Qu'il nous suffise de répéter ici que les prétendues relations constatées entre les deux familles se résument en étymologies futiles, dénuées de tout caractère scientifique. On les donnerait toutes volontiers pour la moindre raison tirée de la forme même des notes.

Deux causes principales semblent avoir présidé à la conception si peu scientifique d'une origine commune des langues indo-européennes et des langues sémitiques.

Nous trouvons la première de ces causes dans la nationalité, ou, pour mieux dire, dans la race même d'un certain nombre des auteurs qui ont soutenu cette opinion. Une grande partie d'entre eux sont israélites. Nous n'avons que faire de citer des noms ; le fait est assez connu. Il y a là un sentiment que nous n'analysons point, que nous n'approuvons pas davantage, mais que cependant nous pouvons comprendre.

La seconde cause c'est l'esprit biblique, l'esprit de secte, pour lequel il ne saurait exister de vérité en dehors de la théologie. Avant tout, il anathématise l'examen libre et laïque, quitte à faire volte-face au dernier moment et à déclarer que toutes les connaissances acquises procèdent de lui et de lui seul. L'esprit biblique a décidé jusqu'à nouvel ordre que l'hébreu et le sanskrit avaient des racines communes. Soit. Nous prenons acte de cette décision, mais sans lui donner plus de valeur. L'on ne discute point avec des gens qui se proclament eux-mêmes illuminés. Au surplus on comprend assez quelle sorte d'intérêt pousse ces derniers tenants de la sainte Ecriture à assigner à toutes les langues de l'univers une origine commune, et combien il leur importe, en particulier, de les rattacher plus ou moins directement à la prétendue langue du premier des Juifs. Laissons, comme dit la même Ecriture, laissons les morts enterrer leurs morts.

Est-il possible de déterminer d'une façon précise la région où fut parlée la langue sémitique commune, l'idiome d'où procédèrent l'araméen, l'assyrien, l'hébreu, l'arabe ?

Cette question, à notre sens, est assez difficile à résoudre. On n'a pas laissé que de l'aborder. F. Schrader divise la famille sémitique en deux groupes : un groupe du nord comprenant l'araméen, l'assyrien, l'hébreu et le phénicien ; un groupe méridional comprenant les deux subdivisions du même groupe que nous avons indiquées

ci-dessus. Ces deux groupes se distinguent nettement l'un de l'autre et par leur mythologie et par leur langue. Nous savons par les traditions de l'antiquité que les Phéniciens venaient de Babylonie, les Hébreux de Mésopotamie et de Babylonie. C'est ce qu'enseigne également toute leur civilisation, et la linguistique rapproche aussi les Hébreux et les Phéniciens des Assyriens, sans que pourtant la langue assyrienne soit la source même des langues du groupe chananéen. Quant au groupe méridional, le groupe arabe, il ne peut manifestement provenir du précédent. Sa mythologie l'en diversifie tout à fait et sa langue est incontestablement plus pure, plus rapprochée du type sémitique commun que ne l'est l'assyrien, l'araméen ou l'hébreu.

En somme, ce serait dans l'Arabie du nord ou dans l'Arabie centrale qu'aurait été parlée la langue sémitique commune.

F. Schrader suppose, d'autre part, que la division du système sémitique commun ne s'opéra pas tout d'un coup. Le groupe du nord se sépara le premier, laissant un autre groupe compacte qui se créa alors la forme des pluriels brisés. Une fois séparé du groupe arabe, le groupe septentrional se divisa à son tour. Les Araméens se seraient détachés les premiers à l'époque où ce dernier groupe était encore en Babylonie et ils se seraient les premiers dirigés vers l'ouest (1).

Ce ne sont là, disons-le, que des hypothèses. Peut-être se vérifieront-elles, peut-être leur sera-t-il substitué d'autres suppositions. Elles portent, à la vérité, un certain caractère de vraisemblance, mais nous ne voulons point nous prononcer sur leur compte. Nous pensons que cette question demeurera obscure longtemps encore et que la

(1) *Die Abstammung der Chaldæer und die Ursitze der Semiten.* Zeitschrift der deutschen morgenlœnd. Gesellschaft, t. XXXVII. Leipzig, 1873.

linguistique seule ne la résoudra pas sans le secours de l'anthropologie et de l'archéologie.

B. LES LANGUES KHAMITIQUES.

Il est à peine besoin de dire que ce terme de langues *khamitiques* est tout aussi défectueux que celui des langues *sémitiques*. L'usage pourtant paraît le consacrer, et nous avons dû l'adopter à défaut de toute autre dénomination acceptable.

On a bien proposé le nom de « libyen », mais ce nom dit trop peu, et ne s'aplique qu'à l'une des divisions de la famille khamitique.

Les langues khamitiques qui ont couvert la plus grande partie de l'Egypte et toute la rive africaine de la Méditerranée, ont-elles occupé, à un moment donné, les régions de l'Euphrate et du Tigre (au moins en partie) et ont-elles gagné l'Afrique du nord par la Syrie, la Palestine et l'Arabie Pétrée, c'est ce qu'il est difficile d'assurer, bien que certaines présomptions soient en faveur de cette hypothèse.

On sait encore moins, s'il est possible, dans quelle contrée les langues khamitiques se séparèrent des langues sémitiques. A ce sujet on ne peut affirmer qu'une chose : c'est que cette séparation doit être reportée à une très haute antiquité, à un âge qu'aucune chronologie ne peut nous indiquer. Durant toute la période historique, les langues sémitiques ont fort peu varié, elles ont persisté étonnamment dans leurs anciennes formes : ce fait nous dit déjà que l'époque est bien éloignée où langues sémitiques et langues khamitiques n'étaient pas encore nées, mais où il existait un idiome à jamais perdu dont elles devaient procéder les unes et les autres.

Comme l'a très justement remarqué Fr. Müller, la

parenté des idiomes khamitiques et des idiomes sémitiques est plutôt dans l'identité de l'organisme que dans la coïncidence des formes toutes faites. Les deux familles ont dû se séparer à une époque où leur langue commune était encore dans une période fort peu avancée de développement. De plus, le groupe khamitique semble s'être divisé de très bonne heure en différents idiomes et les langues de ce groupe sont bien moins rapprochées les unes des autres que ne le sont entre elles les langues sémitiques.

Le système pronominal des deux familles a tout particulièrement servi à établir leur parenté. Il y a identité de racines entre les pronoms sémitiques et les pronoms khamitiques (Maspéro, *Des pronoms personnels en égyptien et dans les langues sémitiques*, 1873), et nombre d'autres liens de parenté ont été mis en évidence d'une façon indiscutable.

Un élément *t* forme le féminin, que l'on tire de la forme masculine : égyptien, *son* « frère », *sont* « sœur » ; bédja, *tak* « homme », *takat* « femme » ; tamachek, *akli* « nègre » *taklit* « négresse » avec préfixation et suffixation de l'élément dérivatif.

La formation du pluriel a lieu, en égyptien, par le suffixe *u* : *sonu*, *sontu* « frères, sœurs ». Cet *u* est assurément parent du dérivatif *an* que l'on trouve en tamachek.

En général, on ne trouve plus de traces de déclinaison, c'est-à-dire de suffixes indiquant les cas. On a communément recours à des particules placées avant ou après le nom, pour exprimer les relations du nom en question avec le reste de la phrase.

Les formes de la conjugaison sont nombreuses, comme dans les langues sémitiques. Quant au système des temps il est très élémentaire, comme le système du sémitisme.

C'est d'ailleurs ce que nous allons voir dans chacun

des paragraphes consacrés ci-dessous aux différents idiomes khamitiques.

Les pronoms possessifs sont tantôt préfixes (en saho, en dankali), tantôt suffixes (égyptien *pera* « ma maison », *perua* « mes maisons »).

La numération est décimale, mais on trouve en bédja et dans quelques autres idiomes méridionaux la trace du système quinaire (Prætorius, *Zeitscher. der deutschen morgent. Gesellsch.*, tome XXIV, p. 415).

On distingue trois groupes dans la famille khamitique : le groupe ÉGYPTIEN, le groupe LIBIEN, le groupe ÉTHIOPIEN. Ils vont nous occuper successivement, au moins d'une façon rapide.

I. *Groupe égyptien.*

C'est au commencement de ce siècle que les anciens hiéroglyphes d'Egypte furent déchiffrés. Depuis bien des centaines d'années ils n'étaient plus que lettre morte. Leur déchiffrement a illustré le nom de Champollion, qui, s'il ne fut point le seul interprète de ces textes précieux, a fait incontestablement plus que tout autre pour les progrès de leur lecture.

Un mot d'abord sur les hiéroglyphes (1).

Le nombre de ces caractères est considérable ; il y en a de phonétiques, il y en a de figuratifs. Les signes phonétiques se transcrivent aisément en caractères latins. Souvent les Egyptiens n'ont écrit d'un mot que les consonnes et ont négligé les voyelles ; mais d'ordinaire il est facile de rétablir ces dernières, soit d'après le sens même du reste de la phrase, soit en comparant le mot en question avec le mot correspondant de la langue copte, dont nous allons avoir à parler tout à l'heure.

Ajoutons que les signes phonétiques peuvent être sim-

(1) BRUGSCH, *Grammaire hiéroglyphique*. Leipzig, 1872.

plement alphabétiques, c'est-à-dire n'exprimer, par exemple, qu'une seule consonne, ou bien syllabiques, c'est-à-dire rendre toute une syllabe. Naturellement leur transcription est aussi aisée dans les deux cas. Quant aux signes figuratifs, ce sont de vrais et purs dessins. Ils se trouvent placés à la fin des mots écrits en caractères phonétiques et ont pour mission de déterminer avec plus de précision le sens même de ces mots. Parfois l'on ne rencontre dans un texte que des signes figuratifs. La difficulté alors est grande pour le lecteur, qui doit, en ce cas, se reporter aux variantes de ce texte.

L'égyptien connaissait un article défini et un article indéfini (*le* chien, *un* chien). Il possédait les deux genres masculin et féminin. L'élément caractéristique de ce dernier est *t*. Ainsi le mot *nofer* « jeune homme » a pour féminin *nofert* « jeune fille » ; *son* veut dire « frère », *sont* veut dire « sœur ». Notons que ce *t* peut être aussi bien placé avant le mot qu'après.

La terminaison du duel est *ui* pour le masculin, *ti* pour le féminin : *sonui* « deux frères ».

Le signe du pluriel est *u* pour les deux genres : *sonu* « frères », *tefu* « pères », de *son* « frère », *tef* « père ».

Quant à la déclinaison proprement dite, il n'en existe plus de traces.

L'adjectif, en principe, suit immédiatement le nom auquel il sert d'épithète et il s'accorde avec lui en genre et en nombre. L'on dit, par exemple : *sat urt* « fille aînée, fille grande », *âmu uru* « grands maîtres ». Dans le premier de ces exemples *t* indique un singulier féminin, dans le second *u* indique le pluriel.

Le pronom possessif est suffixé au nom : première personne du singulier *-a* ; deuxième personne sing. masc. *-k*, féminin *-t* ; troisième personne masc. *f*, féminin *s*, etc. Exemples : *pera* « ma maison », *perk* ou *pert* « ta maison », etc. ; au pluriel *perua, peruk, perut*, etc.

Le sujet dans la phrase égyptienne se place parfois avant le verbe, mais pour l'ordinaire le verbe occupe la première place, le sujet vient ensuite, puis le régime direct, puis le régime indirect, puis l'adverbe.

Dans les formes verbales l'élément personnel se place à la fin du radical, à la fin du thème :

uonk « tu es », masculin ;
uont « tu es », féminin ;
uonf « il est » ;
uons « elle est » ;
uonten « vous êtes » ;
uonu « ils sont ».

Comme on le voit, le procédé de formation est absolument comparable au procédé de formation du nom possessif.

Nous n'avons parlé, en commençant ce paragraphe, que de l'*écriture hiéroglyphe*. Il est aisé de concevoir que ce système dut se simplifier dans la suite des temps et se modifier en une large mesure pour répondre aux besoins des relations ordinaires de la vie. Il donna naissance à deux écritures cursives, l'une appelée *hiératique*, l'autre appelée *démotique*. Au livre second de ses Histoires, Hérodote parle de la double écriture égyptienne, l'une sacrée, l'autre populaire. L'écriture hiératique, écrit de droite à gauche, reproduit simplement sous une forme cursive et souvent très abrégée les anciens hiéroglyphes. On la trouve rarement sur les monuments de pierre ; le plus souvent on la rencontre sur des papyrus. C'était l'écriture savante et religieuse.

L'écriture démotique n'est, à son tour, qu'une forme de l'écriture hiératique et elle contient encore un certain nombre de véritables idéogrammes. C'était l'écriture populaire, qui servait à transcrire la langue courante, la langue vulgaire (1), cette langue qui peut expliquer bien

(1) Brugsch, *Grammaire démotique*. Berlin, 1855.

des différences existant entre l'égyptien ancien et le copte.

L'écriture démotique se lit de droite à gauche comme l'écriture hiératique. Assez rapprochée tout d'abord de cette dernière, elle finit par s'en distinguer très nettement et devint à l'époque de la domination romaine fine et dégagée.

La littérature écrite en caractères démotiques comprend des documents officiels, tels que les décrets ; des inscriptions dédicatoires gravées sur pierre ; des contrats de vente sur papyrus, des inscriptions funéraires.

C'est au musée de Turin que se trouvent les plus anciens papyrus.

Le démotique littéraire comprend une période de dix siècles environ ; dans la première partie du troisième siècle de notre ère, il était encore en usage.

La langue *copte* procède de l'ancien égyptien, sa période littéraire va du troisième au septième siècle de notre ère. C'est une littérature toute chrétienne et qui est assez considérable. L'islamisme ruina la langue copte et lui substitua l'arabe partout où elle servait d'idiome populaire. Elle continua pourtant de mener dans quelques monastères une existence à peu près factice, mais aujourd'hui elle est complètement éteinte.

Le matériel phonétique du copte était plus riche que celui de l'ancien égyptien, mais sa grammaire n'en différait pas d'une façon notable. A qui connaît la langue copte, il est donc facile d'apprendre l'ancien égyptien, et réciproquement. Le vocabulaire copte comprend toutefois un assez grand nombre de mots empruntés au grec.

Le copte marque le féminin en préfixant au nom l'élément *te*, *ti*, qui n'est autre que l'ancien article féminin. Il peut encore recourir à l'allongement de la voyelle terminale (*uro* « roi» , *urô* « reine ») ou à un *i* suffixé. Le

signe du pluriel est *u*, comme en ancien égyptien, et de plus il en est un second, *i*, qui peut se combiner avec le premier : *sbô* « enseignement », au pluriel *sbôui*. Quant aux cas, il n'y en a point de traces : ce sont des particules placées devant le nom qui en expriment l'idée.

Le verbe copte possède la double formation par préfixes et par suffixes, qu'il est aisé de comparer à la double formation sémitique dont nous avons parlé plus haut, en traitant du sémitisme en général. Ainsi le signe *k* « tu », de la seconde personne du genre masculin, est placé parfois avant le thème verbal, parfois après. Mais cette différence de position n'est point capitale comme dans le verbe sémitique (où elle donne à entendre que l'action est accomplie ou qu'elle ne l'est pas) ; ici, au contraire, que l'élément personnel soit posé avant ou après le thème verbal, cela ne paraît pas avoir d'influence sur le sens du mot. C'est au moyen de verbes auxiliaires préposés au thème verbal que le copte distingue les différents temps, son parfait, son futur, etc.

L'alphabet copte n'est autre que l'alphabet grec, dont les caractères sont un peu plus gras et plus arrondis ; parfois on les incline légèrement vers la gauche. Quelques signes supplémentaires ont été ajoutés d'ailleurs à l'alphabet grec pour rendre les sons particuliers au copte que le grec ne possédait pas, par exemple notre « ch ».

On distingue en copte trois dialectes : celui de Memphis, qui possédait les aspirées *kh*, *th*, *ph* ; celui de Thèbes au sud, et un dialecte du nord.

II. *Groupe libyen.*

L'ancien libyen occupait le nord de l'Afrique, à l'ouest de l'égyptien, et c'est sur son domaine que vint s'implanter le punique, le phénicien d'Afrique. La grammaire de l'ancien libyen n'est pas encore rédigée, mais on commence à la connaître par ses inscriptions. Faidherbe en a publié une très importante collection, au nombre de deux cents environ, dont plusieurs sont bilingues, une accompagnée d'un texte phénicien, d'autres accompagnées d'un texte latin (1).

Le libyen actuel n'a pas un nom général dont l'emploi soit adopté communément. Le nom de *berber* ou berbère est peut-être appelé à devenir une dénomination commune ; quant à ceux de « kabyle », de « tamachek » et à bien d'autres encore, ce ne sont que des applications particulières à tel ou tel dialecte et que l'on ne peut étendre à leur ensemble.

Il est difficile de dresser exactement la carte de la langue berbère. Toute la partie méridionale de Tripoli, de la Tunisie, de l'Algérie, du Maroc semble lui appartenir, et, en certains endroits, elle longe encore la Méditerranée, par exemple, en Algérie, de Dellys à Bougie et au delà encore vers l'est (Kabylie), entre Tenès et Cherchell (2).

Comme dans les autres langues khamitiques, un *t* est le signe du féminin. Parfois ce *t* ne figure qu'au commencement du mot, par exemple, dans *tes* « vaches » et dans *tamaher* « femme touarègue » féminin de *amaher* « toua-

(1) *Collection complète des inscriptions numidiques*, Mémoires de la Société des sciences de Lille, 3e série, t. VIII, p. 361. Paris, Lille, 1870.

(2) Hanoteau, *Essai de grammaire de la langue tamachek*, in fine. Paris, 1860.

rég » ; mais, pour l'ordinaire, il est placé tout à la fois avant et après le mot, ainsi que le montrent les exemples suivants :

akli « nègre », *taklit* « négresse » ;
ekahi « coq », *tekahit* « poule » ;
aluki « veau », *talukit* « génisse » ;
amekkelu « sorcier », *tamekkelut* « sorcière ».

Le préfixe initial n'est que l'ancien article féminin soudé au substantif.

Le pluriel est en *an*, *en*, au féminin *in*. Sa formation demanderait, d'ailleurs, quelques explications un peu étendues : *amenukal* « roi » fait, par exemple, *imenukalen*.

La déclinaison est remplacée ici aussi par l'emploi de prépositions : *ales en tamet'* « l'homme de la femme, le mari de la femme », *ifka i ales* « il donna à l'homme ».

Quant au verbe, le berber ne possède qu'une seule forme, une espèce de forme aoristique, indéterminée, à laquelle on prête l'idée de présent ou de futur par des procédés tout à fait accessoires. L'élément personnel est placé après le thème verbal, sauf à la première personne du singulier et à la troisième personne du pluriel. Ainsi le verbe *elkem* « suivre » fait *telkem* « elle suit » et *elkemenet* « elles suivent » (*t-elkem*, *elkem-en-et*).

Nombre de mots arabes se sont glissés dans les différents dialectes berbers. Ceux-ci ont, d'ailleurs, perdu toute espèce d'écriture propre, sauf le dialecte tamachek.

Cette écriture formée de signes assez réguliers, est difficile à lire ; les voyelles n'y sont pas représentées et les mots ne sont point séparés les uns des autres. Pour la déchiffrer, il faut donc, avant tout, connaître la langue elle-même qu'elle représente.

Hanoteau compte en Algérie plus de 855.000 Berbers. Le département d'Oran n'en contient qu'un petit nombre. Ils occupent plus particulièrement les deux autres

départements ; dans celui de Constantine il y en aurait près de 500.000 Quant aux Berbers que l'on rencontre au sud de l'Algérie, il est fort difficile de savoir à quel nombre ils s'élèvent.

Ajoutons que la langue des Guanches, anciens habitants des Canaries, se rattachait au groupe libyen (1).

On trouve en tamachek des témoignages évidents de flexion. Il suffit de comparer les formes *selmed* « enseigner », *seggech* « faire entrer », *toukemmet* « être cueilli », et les formes *salmad* « avoir l'habitude d'enseigner », *seggach* « avoir l'habitude de faire entrer », *toukemmat* « être habituellement cueilli ». Un changement vocalique sert à former du temps aoriste le présent défini : *telkemed* « tu as suivi, tu suivais, tu suivis », *telkamed* « je suis en ce moment » Fr. Müller, *Grundriss*, t. III, p. 283). En kabyle, l'idée d'habitude peut être rendue par un redoublement de consonne (*zer* « voir », *sel* « entendre », d'où *zerr* « voir habituellement », *sell*, etc.), mais aussi par une mutation vocalique : *gen* « dormir », *gan* « dormir habituellement », *senger* « dévaster », *serkem* « faire bouillir », cf. *sengar*, *serkam*. De même *sousef* « cracher », *sousouf* « avoir l'habitude de cracher » ; *souther* « demander », *southour* « avoir l'habitude de demander ».

III. *Groupe éthiopien.*

Les langues qui composent cette famille ne doivent pas être confondues avec les idiomes sémitiques de l'Abyssinie, tigré, amharique et autres, dont nous avons parlé ci-dessus et qui forment une division du groupe arabe. A ces derniers idiomes on a donnné parfois le

(1) Sabin Berthelot, *Mémoires sur les Guanches*. Deuxième partie. Mémoires de la Société ethnologique, t. II, p. 77. Paris, 1845.

nom de langues éthiopiennes. Cela prête à confusion. Le nom de langues éthiopiennes doit être réservé au groupe khamitique des langues de l'Afrique centrale parlées au sud de l'Egypte, aux alentours et dans certaines parties de l'Abyssinie.

On en compte six principales : le *somâli*, répandu dans le territoire en forme de coin qui s'étend au sud du détroit Bab-el-Mandeb, ainsi que du cap Guardafui, et qui donne, à l'est, sur la mer des Indes ; le *galla*, parlé à l'ouest du somâli, dans l'intérieur des terres, au sud de l'Abyssinie et au nord des langues appartenant au système bantou ; le *bedja*, langue des Hadendoas et d'une partie des Beni-Amer (également appelé le bédouié, la langue des Bédouins), parlé entre le Nil et la mer Rouge au nord de l'Abyssinie ; le *saho ;* le *dankâli ;* l'*agaou*.

La classification de ces différents idiomes n'est pas encore établie ; dans l'état actuel des connaissances, on ne peut que les grouper les unes avec les autres et les rattacher aux autres langues khamitiques. Ils en présentent d'une façon très évidente tous les caractères.

Le bédja possède un article, l'article déterminé. C'est, au singulier *û* pour le nominatif masculin, *tû* pour le nominatif féminin ; au pluriel, *â* pour le nominatif masculin, *tâ* pour le nominatif féminin (cf. Almkvist, *Die Bischari-Sprache Tû-Bedawie*) ; on reconnaît l'élément féminin *t*.

Comparez la formation des substantifs tels que *tak* « homme », *takat* « femme ». Parfois le signe du féminin est à la fois préfixé (ancien article) et suffixé.

Tandis que le pronom possessif est suffixé au nom en bédja, en galla, en somâli (comme en égyptien et en tamachek), en dankali, en saho il est préfixé : dankali, *ku-kitâba* « ton livre « ; saho, *ku-inâ* « ta mère ».

Quant à la conjugaison elle-même, nous trouvons en saho (comme en copte) une forme où l'élément personnel

précède le radical et une forme où il le suit : il le précède dans *nekke* « nous étions » et le suit dans *kino* « nous sommes » (*ne-kke*, *ki-no*). Il en est de même du galla, qui dit, par exemple : *gigna* « nous allâmes » et *nefdeg* « nous perdîmes » (*gig-na ne-fdeg*) ; ici la première forme est celle d'un parfait et la seconde est celle d'un temps aoristique, indéterminé. Ce procédé est analogue à celui qu'emploient les langues sémitiques en semblable occurence.

C. LES LANGUES INDO-EUROPÉENNES.

Nous aurons à donner plus de détails sur les langues de cette famille que nous n'en avons donné sur chacune des autres.

La raison en est simple. L'importance des langues indo-européennes est grande à tous les points de vue. Après avoir été les interprètes de la civilisation hindoue, de la civilisation éranienne, de la civilisation grecque et de la civilisaton latine, elles servent d'organe aujourd'hui à la civilisation moderne. Elles étouffent peu à peu, au moins en Occident, les idiomes étrangers qui se trouvent en contact avec elles, le basque, le magyar, bien d'autres encore. Il n'est point de langues qui aient autant vécu, qui aient passé par autant de phases et de périodes successives.

Une autre considération doit aussi nous intéresser particulièrement. Les langues indo-européennes possèdent seules une véritable grammaire comparée. Nous avons dit que la grammaire des langues sémitiques était encore à présent à peu près complète, non seulement dans ses grandes lignes, dans ses traits généraux, mais encore dans une foule de détails.

La découverte du sanskrit, de la langue sacrée des Hindous, devait avancer considérablement la connaissance méthodique des langues européennes appartenant

à la même famille que lui. « Filippo Sasseti, noble marchand florentin, fut peut-être le premier, dans une lettre à Pier Vettori (15 janvier 1585), à parler de la langue en usage alors dans l'Inde, où il se trouvait, langue différente de celle de la religion et de la littérature, et à noter quelques ressemblances entre les noms italiens et indiens (1). » En 1767, le missionnaire Cœurdoux communiquait à l'Académie des Inscriptions ses observations sur la ressemblance d'un grand nombre de mots du latin et du grec avec des mots de la langue « samscroutane ». Vingt ans après, W. Jones proclamait d'une façon définitive le fait de cette parenté. En 1786, il s'exprimait en ces termes devant la Société asiatique de Calcutta : « La langue sanskrite, quelle que soit son antiquité, est d'une admirable structure; plus parfaite que le grec, plus riche que le latin, plus affinée que toutes deux, néanmoins reliée à l'un et à l'autre par une parenté trop grande, tant dans les racines des verbes que dans les formes grammaticales, pour que cela soit purement l'effet du hasard; ces ressemblances sont si frappantes, que nul philologue ne pourrait examiner ces trois langues sans penser qu'elles sont sorties d'une source commune, qui peut-être n'existe plus depuis longtemps. Il y a une raison semblable, bien que moins évidente, pour supposer que le *gothique* et le *celte* ont eu la même origine que le sanskrit; on pourrait aussi adjoindre à la même famille le *persan* antique (2). »

De son côté, Bopp démontra le premier par l'analyse même des formes linguistiques, l'identité de la plus grande partie des langues indo-européennes. Il ne lui fut pas donné de codifier définitivement leurs lois phonétiques, leurs procédés de formation des mots, et sa Gram-

(1) D. Pezzi, *Introduction à l'étude de la science du langage*. Travail par V. Nourrisson. Paris, 1873.
(2) *Op. cit.*, p. 69.

maire comparée n'est plus aujourd'hui qu'une œuvre historique, mais son nom n'en reste pas moins attaché à l'une des découvertes qui font l'honneur du dix-neuvième siècle.

Bopp avait visé dans tous ses écrits à démontrer l'alliance intime du sanskrit, du zend et du perse, du grec, du latin, des langues celtiques, germaniques, slaves et du lithuanien. Cette démonstration une fois acquise, et bien acquise, la science des langues indo-européennes fit un nouveau pas, un pas énorme.

De la parenté de tous ces idiomes, on conclut à une forme antique dont ils seraient tous sortis, forme lointaine, forme à jamais perdue, mais qu'il s'agissait de restituer.

Il n'est que juste de citer ici deux noms que l'histoire de la linguistique ne saurait oublier sans ingratitude, celui de Schleicher et celui de H. Chavée. C'est à ces deux auteurs que l'on doit la première mise en réalisation de cette conception féconde d'une forme commune primitive des langues indo-européennes. Dans l'introduction d'un écrit important, publié il y a déjà quarante ans, H. Chavée pouvait dire : « Ces langues ne sont pour le linguiste que des variétés d'une langue unique et primordiale parlée jadis au centre de l'Asie. Pénétré de cette vérité, nous avons entrepris de reconstituer organiquement les mots de cette langue primitive en rétablissant partout le type original à l'aide de ses variétés les mieux conservées (1). » La linguistique moderne est là tout entière. Schleicher, de son côté, faisait paraître cet admirable manuel que l'on pourra sans doute reviser, compléter, améliorer, mais qui demeurera toujours la base même des études de linguistique indo-européenne (2).

(1) *Lexicologie indo-européenne.* Paris. 1849.
(2) *Compendium der vergleichenden Grammatik der indogermanischen Sprachen*, 3e édit. (posthume). Weimar. 1871.

La langue commune indo-européenne.

Avant de parler des différents idiomes du système indo-européen, avant de rechercher le degré de parenté qui unit de plus près certains d'entre eux, nous avons à esquisser un tableau général de la langue commune qui a donné naissance à ces différents idiomes.

Le tableau que Schleicher en a tracé dans son *Compendium* était le suivant :

Trois voyelles *a*, *i*, *u*, pouvant donner (par la préfixation d'un *a*) les groupes *â* (=*a*+*a*), *ai*, *au*, et, par une seconde préfixation *âi*, *âu* ; les explosives *k*, *t*, *p* ; *b*, *d*, *g* ; *gh*, *dh*, *bh*; les nasales *m*, *n*; *r*; *v*, *y*; la sifflante *s*.

Formation des mots par suffixes : **svap* « dormir », *svapna-* « sommeil » (cf- sanskrit *svapna-*, zend *qaphna-* grec ὕπνο-).

Huit cas : nominatif singulier, * avi-s « le mouton, la brebis », accusatif * avi-m, ablatif * avay-at, génitif *avay-as, locatif* avay-i, datif* avay-ai, premier instrumental avy-â; deuxième instrumental* avi-bhi. Trois nombres ; singulier, duel, pluriel.

En principe c'est par la désinence indicatrice du cas que le genre lui-même est désigné. Ainsi dans les thèmes finissant par un *a* l'élément du cas nominatif, au singulier, est *s*, au neutre cet élément est *m*, le même que celui de l'accusatif. Exemples : AKVA-S « le cheval » (en sanskrit *açvas*, en latin *equus*); YUGA-M « le joug » (en sanskrit *yugam*, en latin *jugum*). Le signe du pluriel suit en principe celui du cas, mais ce signe n'est pas toujours le même et souvent il est fort difficile de découvrir sa forme primitive. En bien des cas c'est simplement la consonne *s*, reste d'un élément qui se montrait jadis dans sa forme intégrale.

Il ne faut pas l'oublier, ces suffixes indiquant le cas et ces autres suffixes indiquant le nombre ont été primitivement des formes indépendantes; ce n'est que par la suite des temps que ces formes en sont arrivées à n'être plus que des éléments secondaires, des éléments destinés à indiquer les relations et les modes d'être d'une autre racine. On a souvent cherché à découvrir la forme primitive de ces éléments ; toutes les tentatives sont demeurées sans résultats certains. L'on a proposé des conjectures plus ou moins probables, mais, en réalité, la solution de ce difficile problème est encore à trouver. Au moins le but auquel il faut tendre est constant, bien établi et vraisemblablement on l'atteindra un jour ou l'autre.

Le verbe indo-européen possède deux voix : l'une transitive « j'entends, je frappe », l'autre intransitive « je m'entends je me frappe », mais toutes deux actives. C'est dans l'élément pronominal placé à la suite du thème verbal qu'il faut chercher l'expression même de cette différence de sens. Il y a, en un mot, deux sortes de suffixes personnels : des suffixes transitifs, des suffixes intransitifs. C'est ainsi, par exemple, qu'à la troisième personne du singulier, le suffixe de la voix transitive est TI et que celui de la voix intransitive est TAI : on reconnaît la forme grecque ται de la voix appelée « passive » par les grammairiens, qui, en effet, à ce sens dans la langue grecque, mais dont le sens premier était simplement intransitif, réflexif. Les suffixes personnels de l'intransitif procèdent, d'ailleurs, des suffixes personnels, de la voix transitive ; celui de la première personne veut évidemment dire « je me », celui de la seconde « tu te », celui de la troisième « il se » (en latin « ego me, tu te, ille se »).

Tandis que le système sémitique ne connaissait que deux temps, l'un exprimant que l'action était accomplie, l'autre qu'elle ne l'était pas encore l'indo-européen

commun en possédait six. Quatre de ces temps étaient simples, les deux autres étaient composés.

Le *présent* a pour forme la plus simple, la racine telle quelle, suivie du suffixe personnel. Parfois la voyelle de la racine a subi cette augmentation dont nous parlions ci-dessus, par exemple la racine I « aller » devient AI : AITI « il va » (sanskrit *éti*, lithuanien *eiti*). Parfois la racine verbale est dérivée; il s'agit de conjuguer une forme complexe, par exemple le thème BHARA- dont l'élément BHAR est radical et dont l'élément A n'est qu'un élément dérivatif. De là le présent BHARATI « il porte ». Quoi qu'il en soit, le présent est toujours un temps simple, qu'il s'agisse de conjuguer la racine elle-même ou un dérivé de la racine.

L'*imparfait* est formé du thème du présent, soit simple, soit dérivé, auquel se préfixe l'augment A ; de plus, les désinences personnelles sont écourtées; *ti* de la troisième personne devient *t* ; *mi* de la première devient *m*. Ainsi le présent BHARATI « il porte » a pour imparfait ABHARAT « il portait ».

L'*aoriste simple* est caractérisé comme l'imparfait, par l'emploi de l'augment et des suffixes personnels écourtés ; il s'en distingue simplement par ce fait qu'il ne tient pas compte de la forme du présent. En grec, par exemple, la racine θε « poser » se dédouble au présent et donne τίθετε « vous posez ; l'imparfait préfixe l'augment ε à cette forme redoublée et fait ἐτίθετε « vous posiez » ; l'aoriste simple ne tient pas compte du redoublement et fait ἔθετε « vous posâtes ».

Le *parfait* a pour caractéristique le redoublement de la racine.

A ces quatre temps simples s'ajoutent, avons-nous dit, deux temps composés. Le *futur* est l'un de ces deux temps. Il est composé de la racine verbale et d'un élément ASYA, SYA, dont le sens premier semble avoir été celui de

« tendre à être » ; de là, par exemple, le sanskrit *dâ syati* « il donnera ». L'*aoriste composé*, que le sanskrit, le zend, les langues slaves et le grec ont conservé — ce dernier sous le nom d'aoriste premier — a pour caractéristique l'élément SA.

Ces six temps sont complétés, dans l'indo-européen commun, par trois modes : l'indicatif, le conjonctif, l'optatif. L'*indicatif* n'a aucune caractéristique : au mode indicatif la forme du temps reste telle quelle. Il en est différemment des deux autres modes. Le *conjonctif* a pour caractéristique un A placé entre le thème et le suffixe personnel : l'indicatif du temps présent étant ASTI « il est », le conjonctif du même temps sera ASATI. On donne parfois à l'*optatif* le nom de potentiel ; ce mode est formé par l'intercalation d'un élément YA, YA, entre le thème verbal et le suffixe personnel écourté : ASYAT « puisse-t-il être » !

La numération est d'ordre décimal.

La systématisation de Schleicher avait l'immense mérite de formuler sous une forme précise l'état des connaissances. Elle ne devait pas être définitive, mais elle fournit une base de discussion bien déterminée et sur laquelle durent tabler toutes les théories qui suivirent.

C'est dans le tableau des voyelles que Schleicher prête le plus à la critique. Il semble démontré aujourd'hui que ce tableau est plus compliqué (Sayce, *Principes de philologie comparée*, p. XIV). Nous avons insisté sur la primordialité d'une voyelle organique *r* (*Mémoire sur la prononciation et la primordialité du r vocal sanskrit*, 1872). On a restitué également des voyelles nasales : voir Brugmann, *Grundriss der vergleichenden Grammatik der indogermanischen Sprachen*, t. Ier, p. 20 (1). Cet auteur, l'un des chefs de l'école dite des « néo-gram-

(1) Joh. SCHMIDT, *Zur Geschichte des indogermanischen Vocalismus*. Weimar, 1875. — F. DE SAUSSURE, *Mém. de la Soc. de*

mairiens », compte treize voyelles ordinaires, brèves ou longues, huit voyelles nasales, quatre roulantes. Il est difficile de supposer que cette restitution généreuse soit de toute exactitude. Brugmann compte, en outre, seize consonnes explosives, quatre nasales, deux roulantes, une aspiration douce, quatre sifflantes.

Ce qui semble assez généralement accepté aujourd'hui, c'est l'inanité de la théorie de « la gradation », de l'augmentation, des voyelles au moyen d'un *a* préfixé (1). D'après Paul Regnaud (*Essais de linguistique évolutionniste*) la vieille théorie du « guna » des grammairiens hindous est en contradiction avec la loi la plus constante du langage, dont le développement a été dirigé par le principe de la moindre action, du moindre effort. En somme, le point de vue de Panini doit être abandonné; il est dépassé.

En somme, il y aurait à revenir absolument sur l'opinion de Bopp, de Schleicher, de Chavée, de Curtius, qui voyaient dans le sanskrit le représentant le plus fidèle du type indo-européen commun. L'école des néo-grammairiens (Osthoff, Brugmann, Louis Havet, etc.) semble donner à la langue grecque une préférence bien accusée. Il ne faudrait pas admettre, toutefois, que les « néo-grammairiens » aient produit dans la science linguistique une révolution totale. Ainsi que l'a fort justement dit Bréal, ils n'ont fait que marcher en avant dans la voie déjà tracée (2).

Lorsque Paul Regnaud — qui veut regarder *a* comme

linguistique, t. III, p. 358 ; *Mémoire sur le système primitif des voyelles dans les langues indo-européennes*. Paris, 1879.

(1) Cf. Victor Henry, *le Muséon*, t. III, p. 484 ; *Journal asiatique*, juillet 1884, p. 37. — Osthoff, *Die neueste Sprachforschung und die Erklærung des indog.* Ablautes, 1866.

(2) Cf. Johannes Schmidt, *Zeitschrift für vergl. Sprachforsch.*, t. XXVIII, p. 303. — Merlo, *Cenni sullo stato presente della grammatica ariana*. Turin, 1885. — V. Henry, *Revue asiatique*, 24 août 1885.

la seule voyelle primitive, ce en quoi nous ne pouvons le suivre — attaque l'opinion des néo-grammairiens sur la multiplicité extraordinaire des voyelles indo-européennes organiques, il réagit peut-être avec raison contre la réaction soulevée par la systématisation très simple de Schleicher. La vérité sera sans doute entre l'opinion de ce dernier et celle de Brugmann. Cet auteur, semble-t-il, restitue un type commun des nuances vocaliques qui n'ont peut-être été que dialectales.

En ce qui concerne les racines, les néo-grammairiens, au lieu de voir dans les différents états du vocalisme des dérivés d'une même racine, des formes présentant ou non le renforcement, supposent (sans rien préjuger sur leur ordre chronologique) plusieurs manières d'être du radical ; la forme faible correspond aux racines pures des grammairiens hindous et de Bopp ; la forme forte correspond à celle que jadis l'on disait modifiée par « gunation » ; *bhr* « porter » serait forme faible, *bher* serait forme forte ; *sru* « couler », *mn* « penser » seraient formes faibles ; *sreu*, *men* seraient formes fortes. « Toutefois, dit Victor Henry, il ne faudrait pas penser que l'indo-européen commun ne possédât que deux degrés de flexion correspondant respectivement à la racine faible et au « guna » ; il en avait trois : l'état normal (grec εἶδος « figure ») ; l'état réduit (ἴδμεν « nous savons ») ; l'état fléchi οἶδα (« je sais »), de la racine *vid* « voir, savoir ». Si ces trois états, ajoute V. Henry, n'apparaissent plus avec netteté en sanskrit, c'est parce que cette langue, confondant en *a* l'*e* et l'*o* primitifs, a englobé dans son « guna » l'état normal et l'état fléchi de la racine : *vêdas* « connaissance » corrélatif de εἶδος, et *vêda* « je sais », corrélatif de οἶδα » (*Dictionn. des sciences anthropologiques*, au mot SANSKRIT).

Dans la notation des racines à l'état dit fort, les néo-grammairiens substituent un *e* (*men*) à l'*a* correspon-

dant du sanskrit (*man*). C'est particulièrement de la coïncidence de *e* en grec et en latin qu'ils tirent leurs conclusions contre le caractère primitif de l'*a* sanskrit qui correspond à cet *e*.

Un des axiomes de la « néo-grammaire » est le caractère absolu, suivant elle, des lois phonétiques, qui ne souffriraient point d'exceptions. En principe, soit, la transformation suit sa voie ; mais, ainsi que le remarque fort justement Paul Regnaud, dans la lutte il arrive que la victoire demeure parfois à l'ancienne forme, qui reste comme le témoin d'un autre âge; cela est conforme à l'enseignement de l'évolution dans la série animale. La coexistence — qui n'est point rare — de deux formes analogues, dont l'une est physiologiquement transformée et dont l'autre offre un état antérieur à cette transformation, est un fait frappant et significatif: l'hypothèse du caractère absolu des lois phonétiques est incompatible avec celle qui attribue une origine commune aux formes que l'on considère habituellement comme étant les variantes les unes des autres, puisque leurs différences mêmes accusent qu'elles ont subi des lois diverses (*Revue de linguistique*, t. XX, p. 37). Cf. Hugo Schuchardt *Ueber die Lautgesetze*, Berlin, 1885 ; *Revue critique*, 1886, n° 15.

Dans son œuvre capitale déjà citée (*Grundriss*, etc.), datée de 1887, Fr. Müller exprime son opinion sur la restitution de l'indo-européen organique. Il n'admet que les voyelles *a*, *i*, *u*, *à*, soit un vocalisme très simple qui se serait, par la suite, richement développé. Outre les consonnes acceptées par Schleicher et Chavée, il admet des gutturales « antérieures », *k̇*, *ġ*, *ġh* (1). Il repousse catégoriquement, d'autre part, la conception des néo-grammairiens » qui, par exemple, dans les racines où

(1) Cf. Ascoli, *Corsi di Glottologia*, t. I, p. 27. — Louis Havet, *Mémoires de la Soc. de linguistique*, t. II, p. 266.

l'élément vocalique est tantôt *ai*, tantôt *i*, regardent *ai* comme organique et voient en *i* une forme écourtée. D'après lui, la racine indo-européenne ne comporte pas plus que la racine sémitique de voyelle fondamentale : de même qu'il faut, en sémitique, parler des racines *k-t-b* « écrire », *f-r-s* « déchirer », de même, en indo-européen, il faut parler de *s* « être », de *p-t* « tomber » (et non de *as*, *pat* selon Bopp et Schleicher, ou *es*, *pet* selon les néo-grammairiens). Les formes *s-p-t* sont les formes courtes ; *as*, *pat* sont les formes allongées, les formes longues, grâce à l'accession de la voyelle *a*, tout comme *yaug*, *raik*, *kar* sont des formes allongées de *yug*, *rik*, *kr*. D'ailleurs la voyelle fondamentale *a* tantôt conserve sa propre forme, tantôt se change en *e*, *o*, *i*, *u*. Les voyelles *a*, *e*, *o* ne peuvent provenir que d'une voyelle fondamentale *a*. La langue organique indo-européenne ne posséda, pour opérer la flexion, que la voyelle *a* : cette voyelle peut s'être changée en *e* dans cette langue même, mais on n'a pas de preuve certaine de cette évolution : c'est après la séparation en indo-éranien et en européen que s'effectuera, dans ce dernier groupe, le passage de *a* en *e*. Comme on le voit, Fr. Müller repousse absolument la théorie des néo-grammairiens sur le vocalisme des racines.

En parlant des différents idiomes de la famille indo-européenne nous ne pourrons indiquer que d'une façon très rapide ce que chacun de ces idiomes a conservé de l'héritage commun, ce qu'il en a perdu, mais on peut connaître déjà, au moins d'une manière assez générale, quelle était la composition et la richesse de cet héritage.

Nous nous sommes servis jusqu'ici pour désigner cette langue reconstituée dont sont sortis les différents idiomes indo-européens de la seule dénomination d'*indo-européen commun*. Bopp avait donné au sanskrit, aux langues éraniennes, aux langues slaves, germaniques,

celtiques, au grec, aux langues de l'Italie le nom de langues *indo-germaniques*. Cette appellation, qui prévaut encore aujourd'hui en Allemagne, ne supporte pas la moindre critique ; à quelque point de vue que l'on veuille se placer, elle est absolument vicieuse. Pourquoi ne pas dire aussi bien langues indo-italiques, langues indo-slaves ? Quelques auteurs ont proposé avec plus de raison le nom de langues « indo-celtiques ». Ils se fondaient en cela sur une sorte de motif géographique et leur idée était assez acceptable. On semble ne l'avoir pas goûtée, et il n'est pas à croire que ce nom ait plus de succès que n'en a eu celui de syro-arabe que l'on a voulu donner au système des langues sémitiques. Indo-celtique, d'ailleurs, n'est pas irréprochable : le second terme du mot est exact, le premier ne l'est point. L'Inde, en effet, n'est pas occupée seulement par les idiomes alliés au sanskrit, elle possède également les langues dravidiennes qui n'ont avec les précédentes aucun lien de parenté.

Un nom plus court et qui a paru un moment devoir faire son chemin a été proposé : celui de langues aryennes. On est parti de ce prétendu fait que les anciens Hindous et les anciens Eraniens se donnaient à eux-mêmes le nom d'Aryas (1); mais il est hors de doute que rien n'est moins prouvé, nous dirons même que rien n'est moins vraisemblable. Peu importe dès lors que la racine que l'on retrouve dans le sanskrit *arya-*, *ârya-* « noble », dans le zend *airya-*, existe aussi dans les autres langues de la même famille, par exemple dans les langues celtiques. Peu importe que le nom d'Arie ait été donné à une région spéciale dont les habitants pouvaient recevoir à juste titre le nom d'Aryens. La question est tout autre. Il s'agit de savoir si ce terme peut être généralisé, s'il est permis de l'étendre à toute la famille.

(1) Pictet, *les Origines indo-européennes ou les Aryas primitifs*, t. I, p. 28. Paris, 1859.

Sans hésiter nous répondrons que cela n'est point justifié. Il n'y a pas même un commencement de preuve. Nous ne suivrons donc pas Oppert, Chavée, etc., lorsqu'ils donnent à la langue commune indo-européenne le nom d'aryaque, de langue aryaque, et nous n'accepterons pas davantage le nom de langues aryennes appliqué par un certain nombre d'auteurs aux différentes langues issues de cet ancien idiome.

Le nom d'indo-européen est sans doute un terme fort vague, un terme de convention dont les deux composants disent l'un et l'autre plus qu'ils ne devraient dire — bien que l'un d'eux, d'autre part, soit insuffisant, puisqu'il laisse à l'écart les langues éraniennes ; — sans doute également il y aurait avantage à substituer à ce mot par trop long un terme plus commode ; mais ce dernier terme, il s'agit de le découvrir, et nous estimons que le nom d'indo-européen ne doit être abandonné pour une autre expression que si cette expression se trouve parfaitement justifiée. Ce n'est point le cas des mots aryaque et aryen, aussi n'employons-nous exclusivement que ceux d'indo-européen et de langues indo-européennes.

La famille des langues indo-européennes se divise en huit grandes branches, en huit grands groupes : groupe *hindou*, groupe *éranien*, groupe *hellénique*, groupe *italique*, groupe *celtique*, groupe *germanique*, groupe *slave*, groupe *lettique*. Nous allons les passer successivement en revue ; nous examinerons leur physionomie particulière, nous dresserons le tableau de leurs propres divisions et nous parlerons, à l'occasion, de leur histoire et de leur littérature. Nous aurons à rechercher également quel est le degré de parenté qui peut unir plus intimement les unes aux autres certaines branches de cette grande famille et nous dirons quelques mots de la contrée où aurait été parlé, selon toute vraisemblance, l'indo-européen commun.

§ 1. Branche hindoue.

Dès la fin du seizième siècle, un Italien, Philippo Sassetti, attirait l'attention sur la vieille langue sacrée des Hindous, le sanskrit, et se plaisait à en rapprocher un certain nombre de mots de sa langue maternelle (1). Deux cents ans plus tard, vers la fin du dix-huitième siècle, le moine Paulinus a Sancto Bartholomæo publiait à Rome la première grammaire du sanskrit qui ait été rédigée en une langue européenne. Quelques années auparavant, les Français Cœurdoux et Barthélemy avaient communiqué à l'Académie leur opinion raisonnée de la parenté du sanskrit avec le grec et le latin. Les travaux d'un grand nombre de savants anglais, parmi lesquels nous devons citer William Jones, Colebrooke, Carey, Wilkins, préparèrent et rendirent possible l'œuvre véritablement fondamentale de Bopp. C'était sur le sanskrit que l'on allait établir l'édifice tout entier de la grammaire comparée indo-européenne ; non que cette antique langue pût être regardée, même dans ses monuments les plus anciens, comme la mère commune des langues éraniennes, du grec, du latin, des langues slaves et des autres idiomes de la même famille, mais bien parce qu'en définitive elle s'éloignait beaucoup moins que chacun d'eux de la langue, aujourd'hui perdue, dont ils descendaient tous les uns et les autres. Le grec, le latin et leurs congénères ne procèdent pas plus du sanskrit, que l'hébreu et l'assyrien ne procèdent de l'arabe. C'est donc sans aucune raison que l'on donne parfois aux langues indo-européennes le noms de langues « sanskritiques ». Souvent, à la vérité, les formes du sanskrit sont plus correctes, mieux conservées que celles des autres langues indo-européennes, mais ces dernières, à leur

(1) *Lettere*, p. 415 et suiv. Florence. 1855.

tour, l'emportent parfois sur le sanskrit et se rapprochent avec plus de fidélité du type commun qui leur a donné naissance. Et ce que nous disons ici s'applique tout aussi bien au sanskrit védique qu'au sanskrit classique; cette distinction est oiseuse lorsqu'il s'agit d'établir le degré de parenté des idiomes indo-européens, et le sanskrit védique n'est en réalité que du sanskrit, c'est-à-dire une simple branche de la souche commune des langues indo-européennes.

Le rameau hindou ne comprend, au fond, qu'une seule espèce d'idiomes, mais les uns sont anciens, les autres sont contemporains et nous les examinerons sous deux paragraphes différents.

I. *Anciennes langues hindoues.*

Le mot *sanskrta-* signifie « parfait, accompli » : la langue sanskrite est donc la langue parfaite, accomplie. Ce nom lui a été donné en opposition du terme *prâkrta-* qui veut dire « naturel » et qui s'applique à la vieille langue vulgaire, ou, pour parler plus exactement, aux différents idiomes de la langue vulgaire. Le sanskrit était la langue sacrée, la langue juridique, la langue littéraire ; le prâkrit était la langue courante, la langue populaire, qui tout d'abord ne fut pas une langue écrite.

Le sanskrit possédait les voyelles *a*, *i*, *u* (prononcez « ou ») et leurs longues, les voyelles linguales *r*, *l* et la longue de la première d'entre elles, un *ê* et un *ô;* enfin les diphthongues *âi* et *âu*. Son système de consonnes était fort riche. Outre les explosives *k*, *t*, *p*, *g*, *d*, *b*, *gh*, *dh*, *bh*, il possédait des explosives chuintantes « tch » et « dj » et des explosives linguo-dentales empruntées, selon quelques auteurs, aux idiomes dravidiens et que l'on transcrit d'habitude par un « t » et un « d » munis

d'un point placé au-dessous (*t*, *d*, etc.). De plus, tandis que la langue commune indo-européenne ne connaissait en fait d'aspirées que les consonnes « gh, dh, bh », le sanskrit possédait à côté de chaque explosive forte sa correspondante aspirée, c'est-à-dire *kh*, *th*, *ph*. Cela lui faisait vingt explosives, dont dix simples et dix aspirées. La langue commune indo-européenne ne connaissait que deux nasales « m » et « n » : le sanskrit en possédait une pour chaque ordre de ses consonnes, une labiale, une linguo-dentale, etc., en tout cinq. Au lieu d'une simple sifflante « s », il en possédait quatre. De plus, il avait une aspirée, *h*, enfin un *y* et un *v*.

L'euphonie de la langue sanskrite est des plus compliquées ; on ne peut la bien connaître que par une longue pratique. Ses règles sont d'une sévérité absolue, et si elles reposent, en général, sur des principes acoustiques parfaitement saisissables, on peut dire qu'elles semblent parfois d'une finesse presque exagérée que nous avons quelque peine à comprendre (1). L'euphonie si délicate des langues slaves est loin d'être aussi fine que celle du sanskrit et l'euphonie du grec et du latin n'est en comparaison, qu'une espèce d'essai très rudimentaire.

Par contre, la formation des mots n'offre point de grandes difficultés. Cela tient à la conservation même de la langue. Les éléments qui concourent à la dérivation des mots demeurent bien autrement reconnaissables en sanskrit qu'ils ne le sont dans toute autre langue congénère, sauf peut-être dans les idiomes éraniens.

On peut dire que la déclinaison est à peu de chose près celle de l'indo-européen commun. La plus grande différence que l'on pourrait relever entre le tableau d'une déclinaison sanskrite et celui de la déclinaison de la forme indo-européenne commune correspondante tien-

(1) Nous avons essayé d'en dresser un tableau aussi peu compliqué que possible : *Euphonie sanskrite*, Paris, 1872.

drait aux variations euphoniques auxquelles le sanskrit ne peut se soustraire. Ce n'est pas à dire cependant que sa déclinaison, en dehors de ce fait, soit absolument organique. Non, sans doute. Ainsi, nous ne rencontrons en sanskrit la véritable forme de l'ablatif singulier que dans les noms dont le thème se termine en *a* : les vieilles formes latines « senatud, navaled » et autres n'ont point d'analogues en sanskrit. Mais, en définitive, ce n'est là qu'une exception et l'on peut dire, d'une façon générale, que la déclinaison sanskrite reflète assez fidèlement celle de l'idiome commun dont elle procède. Elle l'emporte, sans conteste, sur la déclinaison des vieilles langues éraniennes qui pourtant peut passer pour très bien conservée.

Le sanskrit possède les six temps de l'indo-européen commun, présent, imparfait, aoriste simple, parfait, futur, aoriste composé, et il s'est créé de plus un nouveau temps, un « conditionnel ». Ce conditionnel n'est autre que le futur dont les suffixes personnels sont écourtés et auquel l'augment a été préfixé : de *bhôtsyati* « il saura », on forme *abhôtsyat* « il saurait ». Le conditionnel sanskrit est donc au futur ce que l'imparfait est au présent.

L'ancien idiome védique ne diffère du sanskrit classique, du sanskrit des épopées hindoues, que d'une façon peu considérable ; les différences que présentent les deux idiomes ne touchent d'ailleurs en rien au fond même de la langue, à sa constitution, à son mode de formation, et nous ne pourrions nous arrêter sur ce sujet qu'en entrant dans une série de détails inopportuns (1).

L'écriture hindoue, dite écriture « dêvanâgarî » — « écriture divine », — composée d'une cinquantaine de signes simples se lisant de droite à gauche et d'une quantité de signes complexes où se trouvent unis deux, trois signes simples, a tout avantage à être transcrite en

(1) Ad. Regnier, *Etude sur l'idiome des Védas*. Paris, 1855.

caractères latins pourvus de signes diacritiques. Une consonne, en principe, ne se lit jamais seule : elle est toujours suivie de la voyelle *a*, à moins qu'un signe accessoire n'indique que la voyelle suivant cette consonne est autre que la voyelle *a*. Un mot finit-il par une consonne et le mot suivant commence-t-il par une voyelle, l'écriture relie ces deux mots; cette difficulté et quelques autres tout aussi sérieuses rendent l'usage du dêvanâgarî peu pratique.

Les plus anciennes inscriptions de l'alphabet hindou ont été gravées sur des rochers vers le troisième siècle avant notre ère ; l'origine de cette écriture paraît connue aujourd'hui et l'on s'accorde généralement à la rattacher à l'ancien alphabet sémitique dont nous avons parlé ci-dessus (1). L'alphabet hindou ne demeura point cantonné dans l'Inde, où il est employé par presque tous les dialectes modernes sous différentes formes ; l'écriture tibétaine en dérive, de même que celle des Javanais, et il a donné naissance également à un certain nombre d'autres alphabets.

Parmi les idiomes prâkrits, les idiomes populaires qui vivaient à côté de la langue littéraire et sacrée, il en est un qui eut une destinée toute particulière. Ce fut le *pâli*. Cet idiome fut celui de la propagande bouddhiste, ce fut la langue spéciale de cette religion douée, comme l'on sait, d'une merveilleuse force d'extension. Sa littérature fut donc très importante. Le pâli ne serait autre que l'ancienne langue populaire du pays de Magadha, dans l'Inde du nord-est, langue très ancienne, assurément, et qui montre parfois une supériorité notable sur les anciens documents prâkrits conservés dans les drames de la vieille littérature hindoue. Il ne change

(1) A. Weber, *Indische Skizzen*, 125. Berlin, 1857. — Fr. Müller, *Reise der oest. fregatte* Novara, Linguistischer Theil, p. 219. Vienne, 1867.

point, par exemple, les « y » en « dj » — comme nous le verrons faire également aux idiomes néo-hindous ; — il a conservé certaines formes et sa conjugaison témoigne également d'un état moins analytique. La voyelle *r* du sanskrit n'existe plus en pâli, la plupart du temps elle est remplacée par un *a* ; les voyelles longues deviennent brèves en certains cas ; les trois sifflantes se confondent en une seule, *s* ; l'assimilation des consonnes se développe de plus en plus ; à la fin des mots il ne peut y avoir qu'une voyelle simple ou une voyelle nasale. Dans la déclinaison le duel est entièrement perdu, le datif est suppléé par la forme du génitif. Telles sont, entre autres, quelques particularités du pâli.

De toutes les langues indo-européennes il en est bien peu qui puissent comparer leur littérature à la littérature de l'Inde ancienne. La littérature hindoue brilla non seulement par sa richesse, par sa variété, mais aussi par l'excellence d'une grande partie de ses œuvres. Albrecht Weber en a donné une esquisse rapide et très fidèle (1). L'ancienne littérature védique comprend d'abord le Rig-Véda, le Sáma-Véda, les deux collections du Yadjur-Véda et l'Atharva-Véda. Le premier de ces Védas est un recueil de chants, d'hymnes religieux ; le second et le troisième contiennent des prières et des formules à réciter durant les sacrifices ; le quatrième est plus récent que les trois autres, beaucoup plus jeune notamment que le premier. Outre les collections d'hymnes, la littérature védique comprend encore les « brâhmanas », écrits où se trouvent un grand nombre de prescriptions rituelles, de traditions, d'explications religieuses, et les « soûtras », espèces d'appendices aux écrits précédents. La période classique est bien plus variée. Elle brille tout d'abord dans la poésie épique, puis

(1) *Akademische Vorlesungen über indische Literaturgeschichte*, seconde édition, Berlin, 1876.

dans le drame, la poésie lyrique, la fable, le conte, les proverbes ; elle a produit enfin des ouvrages considérables de grammaire, de rhétorique, de philosophie, d'astronomie, de médecine, et nombre d'écrits techniques. Vient ensuite la littérature bouddhiste, dont le pâli, ainsi que nous l'avons dit, fut le principal organe.

II. *Langues néo-hindoues.*

Elles sont parlées par environ cent quarante millions d'individus dans la partie septentrionale de l'Inde, qui comprend approximativement les deux tiers de toute la presqu'île.

Le sanskrit n'est point leur auteur direct ; elles procèdent toutes des anciens idiomes prâkrits, des anciennes langues populaires qui étaient parlées à côté du sanskrit, langue religieuse et littéraire. On dit ordinairement que les langues néo-hindoues se sont formées vers le dixième siècle de notre ère, peut-être quelque temps auparavant. C'est là une façon de parler. On ne peut entendre par ces paroles qu'une seule chose, c'est que la forme qu'elles présentent actuellement peut remonter approximativement à cette époque ; quant à leur véritable antiquité, elle est bien plus haute, puisqu'en définitive elles ne sont que la suite des anciens dialectes populaires, des anciens dialectes prâkrits.

Il existe un assez grand nombre de langues néo-hindoues ; quelques-unes d'entre elles n'ont qu'une littérature pauvre, d'autres, par contre, ont une littérature très développée (1).

On peut citer parmi les principaux dialectes : au nord-ouest, le *sindhi* (Haidarabad), le *panjabi* (Multan, Lahore), le *kachmiri ;* au nord, le *nepalais ;* au nord-est,

(1) Consulter R. N. Cust, *A. Sketch of the modern Language of the East Indies*. Londres, 1878. — R. Hoernle, *A. Grammar of the Eastern Hindu*. Londres, 1880.

le *bengali* (Calcutta, l'*assami ;* au centre, l'*hindi*, occupant un très vaste teritoire (Bikanir, Sirhind, Delhi, Agra, Lucknow, Benares) ; au sud-ouest, le *gujurati* (Baroda, Surat), le *marathi* (Bombay,Goa) ; au sud-est, l'*oriya*. — Le nord de l'île de Ceylan est occupé par un idiome dravidien, le *tamoul ;* le sud est le domaine de l'*élou* ou singhalais. D'après Childers, Cust, Ern. Kuhn, cet idiome est aryen, parent des langues de l'Inde du nord. On sait que Ceylan a été colonisé par les Hindous six siècles avant l'ère chrétienne et que le bouddhisme y fut subséquemment introduit.

On donne le nom d'*hindoui* à un idiome qui eut, au moyen âge des langues de l'Inde, une grande expansion littéraire et que représentent aujourd'hui encore certains dialectes du nord-ouest de la péninsule. On a dit avec juste raison que l'hindi n'était que de l'hindoui modernisé ; quant à l'hindoustani, auquel, avons-nous dit, on donne également le nom d'ourdou « langue du camp », il se forma au onzième siècle de notre ère sous l'influence musulmane. Son vocabulaire est plein de mots arabes et persans, et, à la différence des autres idiomes néo-hindous dont les alphabets proviennent du dêvanâgarî sanskrit, il s'écrit avec les caractères persans, c'est-à-dire avec l'alphabet arabe augmenté de quelques signes.

La littérature des langues prâkrites contemporaines et de l'hindoui du moyen âge est considérable et chaque jour l'hindi et l'hindoustani donnent les preuves d'une activité qui leur assure un long avenir (1).

Le caractère général de la phonétique des langues néo-hindoues est une forte tendance à l'assimilation, à la substitution de l'articulation « dj » à un *y* plus ancien, celle assez fréquente de « r » à « d », la simplification du système classique des sifflantes, le changement assez

(1) Garcin de Tassy, *Histoire de la littérature hindoui et hindoustani*, 2 vol. Paris, 1839-1847.

ordinaire des anciennes explosives aspirées de la langue classique (« kh, gh, th, dh », etc.) en une simple aspiration « h ». Le genre neutre a disparu dans la déclinaison de presque tous les idiomes néo-hindous, et souvent des thèmes se terminant primitivement par une voyelle ont rejeté cette voyelle finale et se sont terminés dès lors par une consonne. Quant au nombre pluriel et quant aux cas, ils sont exprimés par des suffixes particuliers qui témoignent assez de la forme vraiment moderne de ces idiomes et de leur passage d'un ancien état synthétique à un état analytique. Il y a là un phénomène dont nous aurons à nous occuper en temps opportun. La conjugaison, elle aussi, est entrée dans la voie analytique : les formes de l'ancien prâkrit, contemporain du sanskrit classique, sont perdues et l'on emploie simplement des formes de participes du présent ou de participes parfaits.

III. *Dialectes des Tsiganes.*

La langue des Tsiganes n'est qu'un dialecte néo-hindou. Il est difficile de dire quelque chose de précis sur l'époque de l'émigration des Tsiganes et du commencement de leurs incursions en Asie et en Europe ; il semble, toutefois, que leur arrivée dans cette dernière contrée n'ait pas été postérieure de beaucoup au douzième ou au treizième siècle de notre ère.

Le fond même de leur langue est absolument hindou; c'est un prâkrit très gâté, souvent très défiguré. Quant au lexique, il est plein d'éléments étrangers, de mots empruntés aux différentes populations qu'ont traversées les bandes tsiganes ou au milieu desquelles elles ont vécu plus ou moins longtemps.

F. Miklosich s'est fondé précisément sur l'état du lexique des différentes tribus tsiganes pour tâcher d'établir

l'itinéraire de leurs migrations. Les éléments persans et arméniens qui s'y rencontrent sembleraient indiquer un antique séjour dans les régions de l'Asie où se parlaient les langues éraniennes. Lorsqu'ils mirent le pied sur notre continent, les Tsiganes se trouvèrent tout d'abord dans une contrée où dominait la langue grecque ; ce fait résulte de ce que chez tous les Tsiganes d'Europe, sans exception, l'on peut constater des éléments empruntés au grec. Ils se dirigèrent de Grèce vers la Roumanie, la Hongrie, la Bohême et la Moravie, l'Allemagne, la Pologne et la Lithuanie, la Russie, les pays scandinaves, l'Italie, le pays basque, l'Angleterre et l'Ecosse, l'Espagne (1). Nous ne parlons ici que des Tsiganes européens. On a beaucoup moins de renseignements sur les Tsiganes d'Asie et sur la quantité d'éléments étrangers qui ont pu s'introduire dans leur langue.

La langue des Kafirs Syapoches (à l'ouest des Dardous, dans le Kafiristan) est également un idiome prâkrit (Cust, *op. cit.*, p. 32).

§ 2. Branche éranienne.

La connaissance des deux langues les plus anciennes de ce groupe, le zend et le perse, offre peu de difficultés aux indianistes. De toutes les langues indo-européennes, les langues éraniennes, en effet, sont les plus rapprochées du sanskrit. En général leur phonétique est moins compliquée, moins délicate que la phonétique des idiomes hindous ; mais sous bien des rapports cependant elle lui est comparable. On peut même assurer que le zend et le vieux perse, le perse de Darius et de Xerxès, l'emportent parfois sur le sanskrit et se rapprochent davan-

(1) Miklosich, *Ueber die Mundarten und die Wanderungen der Zigeuner Europa's.*

tage de la langue commune qui a donné naissance à tous les idiomes indo-européens.

Le nom d'*éranien*, de langues *éraniennes* est incontestablement plus correct que celui d'*iranien*, de langues *iraniennes* qu'emploient un grand nombre d'auteurs. Il rappelle une forme plus antique et nous pensons, avec F. Spiegel, qu'il est bon de lui accorder la préférence.

La classification des langues éraniennes n'est pas encore établie. Il se peut qu'un très petit nombre d'entre celles de ces langues que nous connaissions ne soient pas parentes les unes des autres en ligne directe. A coup sûr il n'en est point parmi elles qui puisse se vanter d'avoir été la mère commune de toutes les autres ; le vieux perse l'emporte parfois sur le zend, parfois le zend l'emporte sur le vieux perse. La seule classification qui semble admissible lorsqu'il s'agit des langues éraniennes est celle que l'on peut emprunter au temps même durant lequel elles ont été parlées. Ainsi l'on classera au rang des anciennes langues éraniennes le *zend*, le *perse* et l'ancien *arménien ;* au rang des langues éraniennes du moyen âge le *huzvârèche*, le *parsi* et l'*arménien* classique ; au rang, des langues modernes le *persan*, l'*arménien* plus récent, l'*afghan*, le *béloutche*, le *kourde*, l'*ossète*, et quelques autres dialectes. C'est ce même ordre que nous allons suivre.

I. *Zend.*

Vers le milieu du siècle dernier, un Français, Anquetil-Duperron, âgé de vingt-trois ans, s'embarquait pour l'Inde comme simple soldat, ne pouvant entreprendre d'une autre façon le lointain voyage qu'il méditait. Le dessein de cet homme courageux dont la science ne pourra jamais oublier le nom, était d'apprendre sur les lieux mêmes les langues du pays. Trompé dans son espoir de pouvoir étudier le sanskrit à Chandernagor, il

gagna Pondichéry, seul et presque sans ressources, après cent jours de marche ; des rivages du golfe du Bengale, il se dirigea vers ceux du golfe d'Oman, atteignit Mahé, puis remonta jusqu'à Surate. C'est là que gagnant la confiance des prêtres d'une colonie de Parses, il fut initié à la connaissance du zend et du huzvârèche. En 1762, Anquetil-Duperron revint en France, sans fortune, mais avec plus de cent manuscrits. Quarante et quelques années après il mourait à Paris, pauvre et indépendant. Les textes qu'il avait rapportés allaient servir à fonder une science nouvelle.

Le *zend* est la langue dans laquelle a été rédigé le texte antique de l'Avesta, livre sacré du zoroastrisme. Nous n'avons à traiter ici ni de l'individualité de Zoroastre ni du contenu des livres sacrés qui lui sont attribués, à lui ou à ses disciples. Il nous suffira de dire que des livres de l'Avesta il ne nous est parvenu qu'une faible partie : le Vendidad, le Vispered, le Yaçna et un certain nombre de morceaux de dévotion intime, de méditations privées, connus sous le nom de petit Avesta. La traduction qu'en donna Anquetil-Duperron était fort défectueuse; elle avait été faite sur les indications sans critiques des prêtres parses, mais Anquetil, en déposant ses manuscrits à la Bibliothèque royale, avait fourni à ses successeurs l'unique moyen de contrôler, de rectifier et de poursuivre son œuvre. Cette tâche échut à un Français, Eugène Burnouf, dont les travaux sur l'ancien perse, langue sœur du zend, ont également illustré le nom. Eugène Burnouf ne fut pas seulement le véritable fondateur de la grammaire zende, il fut aussi le chef de l'école de l'interprétation des livres zends par la tradition, école dont F. Spiegel est aujourd'hui le premier représentant.

Il semble avéré que le zend était la langue des pays éraniens de l'Est, Eugène Burnouf lui donnait pour li-

mites, au nord la Sogdiane, au nord-ouest l'Hyrcanie, au sud l'Arachosie. C'est en adoptant cette opinion qu'on a pu donner au zend le nom de « baktrien », qui, assurément, est fort admissible. Quant au nom de « zend », appliqué à la langue même dans laquelle sont rédigés les anciens textes de l'Avesta, ce n'est qu'une appellation conventionnelle, dont le premier sens n'est peut-être pas encore tout à fait éclairci, mais qu'il serait difficile de détourner aujourd'hui du sens nouveau.

L'alphabet zend est purement alphabétique, c'est-à-dire que chacun de ses signes figure une voyelle ou une consonne. Il contient fort peu de ligatures, et sa lecture, qui se fait de droite à gauche, est assez facile. Son origine est assurément sémitique, mais il ne paraît pas fort ancien et l'on ignore quel pouvait être l'alphabet dont se servaient les Eraniens de l'Est au temps où les Perses, leurs voisins du côté de l'Occident, employaient des caractères cunéiformes.

Le zend comprend deux dialectes, le dialecte ordinaire et le dialecte des Gâthâs. On appelle ainsi un certain nombre de morceaux du Yaçna dont l'interprétation est aujourd'hui des plus difficiles. Les deux dialectes sont fort rapprochés l'un de l'autre ; celui des Gâthâs est regardé d'habitude comme le plus ancien et l'on pense qu'il a été parlé dans les régions montagneuses du pays. C'est encore là une question à éclaircir.

Le systèmes des voyelles zendes est peu compliqué. Outre les *a*, *i*, *u* et leurs correspondantes longues, on y rencontre un *è* long, un *e* qui semble avoir été fort bref, deux autres *e* et deux *o* dont la quantité est variable, un *a* nasalisé est un *â* très labial. Nous avons dit déjà que le zend, plus pur en cela que le sanskrit, n'avait point réduit à une seule et unique voyelle les anciennes diphthongues « ai, au » de la langue commune indo-européenne. Il représente la première par *aê*, la se-

conde par *ao ;* le perse, plus pur encore, conserve les deux diphthongues primitives. C'est ainsi, par exemple, qu'une forme commune *aitat* « cela » donne naissance au sanskrit *êtat*, au zend *aêtat*, au perse *aita*.

Si nous passons aux consonnes du zend, nous avons à remarquer que les sifflantes permutent facilement les unes avec les autres ; c'est là, d'ailleurs, un phénomène commun à l'ensemble des langues éraniennes. Par contre, l'assimilation des consonnes d'ordres divers est assez restreinte, à la différence de ce qui se passe au sanskrit.

En somme la déclinaison du zend est fort bien conservée. Elle garde au singulier, avons-nous dit, l'ancien ablatif en « at », qui a eu un sort si malheureux dans presque toutes les autres langues indo-européennes. De la conjugaison zende nous ne pouvons également dire qu'une chose, c'est qu'elle aussi est bien conservée et qu'elle reflète assez fidèlement le système primitif qui lui a donné naissance.

La question de l'âge de la langue zende peut être résolue, pensons-nous, d'une façon assez approximative. Il est difficile, sans doute, d'émettre une opinion quelconque sur sa première et lointaine origine, on ne peut même se prononcer sur l'époque de sa disparition, mais il est permis de supposer qu'à un moment donné le baktrien fut contemporain de l'ancien perse. A la vérité, nous ne connaissons ce dernier que par les monuments achéménides des sixième, cinquième et quatrième siècles avant notre ère, et il se peut, il est même vraisemblable, que depuis fort longtemps il était déjà parlé. La langue de l'Avesta, le contenu même de ses différents textes, ne permettent pas de l'éloigner considérablement des derniers monuments perses, et à un moment donné les deux langues ont dû être parlées en même temps, l'une, le zend, dans l'Èran oriental, l'autre le perse, dans l'Èran occidental.

II. *Perse.*

Les inscriptions trilingues en caractères cunéiformes, découvertes en Perse sur les ruines des anciens palais et sur le flanc des rochers, étaient rédigées en perse, en assyrien et en une troisième langue dont on ne connaît aujourd'hui encore que fort peu de chose. Nous avons parlé plus haut des diverses tentatives d'interprétation auxquelles a donné lieu ce dernier idiome, dont le texte occupait la seconde colonne (p. 191), et nous avons vu que l'assyrien, langue de la troisième, était un idiome sémitique.

C'est au commencement du XIXe siècle, en 1802, qu'un savant hanovrien, Grotefend, entama le déchiffrement de la première colonne rédigée en perse, où, comme l'on dit souvent, en ancien perse. Le point de départ de sa tentative était ingénieux et simple. Partant de cette idée que des inscriptions, dont quelques-unes avaient dû coûter une peine considérable, relataient naturellement des événements historiques et ne pouvaient être que des annales royales, il remarqua d'abord la fréquence d'un certain groupe de caractères auxquels il attribua le sens de « roi ». Ce groupe était suivi souvent du même groupe augmenté de quelques autres signes ; Grotefend en conclut que ce dernier n'était que le génitif pluriel du premier, et il pensa que les deux groupes, que les deux mots, avaient le sens de « roi des rois ». Le nom qui précédait ces deux groupes devait être forcément un nom propre, et la répétition des mêmes groupes laissait assez entendre qu'il s'agissait d'une indication de généalogies : « Un tel, roi des rois, fils d'un tel roi ». Les recherches de Grotefend furent le point de départ du chiffrement des inscriptions perses. Le Danois Rask les poussa un peu plus loin, mais il était réservé à Eugène Burnouf et à Christian Lassen de donner une véritable lecture de

ces inscriptions, de les interpréter et de formuler leur grammaire. Leurs travaux parurent simultanément, en France et en Allemagne, dans le courant de 1846. La grammaire de l'ancien perse était définitivement connue, elle était comparée méthodiquement à celle du zend et du sanskrit ; la voie, enfin, était ouverte à ceux qui devaient amener l'étude de la langue perse au point où nous la voyons aujourd'hui : Rawlinson, Spiegel (1), Oppert, Kossowicz.

Les inscriptions des rois achéménides n'ont qu'un lexique très restreint, quatre cents et quelques mots environ, y compris un grand nombre de noms propres. Cela est suffisant, toutefois, pour le grammairien. La phonétique du vieux perse, sa déclinaison, sa conjugaison n'ont plus de secrets. Quelques auteurs ont voulu voir dans le perse une forme linguistique plus ancienne que celle du zend ; d'autres auteurs, au contraire, ont prétendu que le zend se rapprochait plus que le perse de la langue commune indo-européenne. On peut défendre, nous semble-t-il, une troisième opinion : le perse, comme nous l'avons dit, l'emporte parfois sur le zend, et parfois le zend l'emporte sur le perse. Tous deux, en principe, ont changé en *h* la sifflante *s* de l'indo-européen commun ; mais le perse, moins correct que le zend, laisse souvent tomber cette aspiration là où ce dernier la conserve. Tandis par exemple, que le sanskrit dit *asmi* « je suis », et le lithuanien *esmi*, le zend dit *ahmi* et le perse *amiy*. Tandis, par contre, que le vieux perse conserve telles quelles les diphtongues *aī*, *au* de l'indo-européen commun, le zend les change en *aê* ou *ôi* et en *ao*. Tous deux, comme on voit, peuvent à juste titre revendiquer la priorité dans un de ces exemples. Il serait aisément superflu, de les multiplier.

Les caractères cunéiformes de la première colonne des

(1) *Die altpersischen keilinschriften*. Leipzig, 1862.

inscriptions trilingues sont loin d'être aussi nombreux que ceux des deux autres colonnes. On en compte environ une soixantaine, tous alphabétiques, c'est-à-dire représentant, non point une syllabe, mais bien une consonne ou une voyelle (1). Ce qui augmente singulièrement leur nombre, c'est ce fait que parfois telle ou telle consonne est représentée par un signe différent selon qu'elle précède telle ou telle voyelle. Entre chaque mot se trouve placé un coin penché obliquement ; on conçoit combien la présence de ce signe a facilité la lecture des textes perses.

Quant à l'origine même de cet alphabet, elle n'est pas encore clairement démontrée ; mais on peut dire, cependant, que les cunéiformes perses ne sont qu'une branche particulière du système général des écritures de cette espèce. Très certainement, c'est la plus simple, ou, pour mieux dire, la plus simplifiée, de toutes ces écritures.

III. *Arménien.*

Il semble que l'arménien se soit distingué de très bonne heure du reste des langues éraniennes ; il occupe en tout cas une place assez indépendante dans la famille des langues éraniennes.

Certains auteurs vont même jusqu'à lui reconnaître une position à part. Hübschmann le regarde comme un idiome indo-européen indépendant de ses congénères (*Zeitschrift* de Kuhn, tome XXIII, p. 407), pourvu de nombreux mots empruntés au persan. Frédéric Müller, dans son plus récent groupement des langues indo-européennes, classe définitivement l'hindou et l'éranien d'une part, l'européen et l'arménien d'autre part, sous

(1) Oppert, *Sur la formation de l'alphabet perse.* Journal asiatique, 1874, p. 238. Paris.

un certain rapport. C'est ce que nous verrons un peu plus loin.

De l'ancienne période de l'arménien nous ne savons que fort peu de chose, particulièrement ce que les auteurs classiques nous en ont transmis. Cette première période prit fin au commencement du cinquième siècle de notre ère. La période de l'arménien classique commence à cette époque. Mesrob crée alors l'alphabet arménien, qui procéderait avec l'alphabet géorgien, comme le dit F. Müller (1), d'une forme sémitique, notamment de l'écriture araméenne. L'âge d'or de l'arménien dura sept cents ans environ et ne prit fin qu'au commencement du douzième siècle. Sa littérature fut féconde, ses dialectes assez nombreux, et l'un d'eux, celui de la province d'Ararat s'éleva bientôt à l'état de langue littéraire. Aujourd'hui encore, les dialectes arméniens sont nombreux, et ce serait une erreur que de les regarder comme des patois de la langue littéraire. Celle-ci, tout au contraire, a paru se fixer et les dialectes actuels ne sont que des formes plus modernes des anciens dialectes. Dès le onzième siècle, on les employait déjà dans la langue écrite, aux dépens de l'idiome littéraire. Ces dialectes semblent se diviser aujourd'hui en deux groupes assez distincts, le groupe oriental comprenant les dialectes d'Arménie, de Géorgie, de la Russie du sud-est, de la Perse, de l'Inde, et le groupe occidental comprenant ceux de Hongrie, de Pologne et de Crimée.

Une des grandes caractéristiques de l'arménien moderne, au moins de l'arménien occidental, est le changement des anciennes explosives fortes en explosives faibles et des anciennes explosives faibles en explosives fortes : les *k*, *t*, *p* anciens deviennent *g*, *d*, *b* et les *g*, *d*, *b* deviennent *k*, *t*, *p*. Le système des consonnes et des voyelles est assez développé. Outre les explosives que

(1) *Ueber den Ursprung der armenischen Schrift*. Vienne, 1865.

nous venons de citer, il comprend un assez grand nombre de sifflantes et deux sortes de *r*.

Par sa déclinaison l'arménien l'emporte de beaucoup sur le persan, dont nous aurons à nous occuper tout à l'heure, et il possède encore, en partie, les anciens éléments indicateurs du cas. Il l'emporte de même, et de beaucoup, dans la conjugaison ; il a conservé, en effet, les anciens temps moins le parfait et s'est créé trois temps nouveaux (un parfait, un plus-que-parfait et un futur) en employant des formes principales. Au demeurant, donc, parmi tous les idiomes néo-éraniens actuellement parlés, l'arménien est celui qui, par la conservation relative de ses formes, se rapproche le plus du type commun de toute la famille. Quant à son lexique, il contient, comme celui de toutes les langues éraniennes modernes, un nombre assez notable de mots étrangers. Les uns ont été tirés du grec au moyen âge ; les autres, plus nombreux, ont été empruntés jadis à l'araméen.

L'arménien a été écrit jadis, sinon d'une façon constante, au moins dans certains documents, en caractères cunéiformes. On a trouvé des inscriptions de cette espèce dans les ruines d'Armavir notamment, non loin du mont Ararat. L'écriture cunéiforme arménienne n'est pas alphabétique comme celle du perse, elle est syllabique ; chaque signe représente, non pas une consonne, ou une voyelle, mais bien une syllabe.

IV. *Huzvârèche.*

L'Avesta, ou, pour mieux dire, les livres de l'Avesta qui subsistaient encore au moyen âge, furent traduits à cette époque dans une langue que nous connaissons non seulement par cette traduction, mais encore par un certain nombre d'inscriptions monétaires et un écrit cosmogogique très important, le Bundehèche. On a tout

d'abord donné à cette langue le nom de *pehlvi*, qui semble un peu trop général ; celui de *huzvârèche*, comme l'ont démontré Joseph Müller et F. Spiegel (1), est le seul qui lui convienne, le seul qu'elle ait porté. On admet assez communément aujourd'hui que cette langue était parlée dans le district occidental de Sevâd. De son origine l'on ne peut rien dire de précis, mais les monnaies dont les légendes sont rédigées en cet idiome et qui datent de l'époque des rois Sassanides nous indiquent qu'elle était encore en usage au milieu du septième siècle de notre ère.

D'après Harlez, le pehlvi est la langue de la version de l'Avesta, et huzvârèche désigne l'ensemble des mots sémitiques qui s'y rencontrent ; pour Sachau, la huzvârèche est le vrai zend.

Parmi les langues sur lesquelles s'est fait sentir le plus fortement l'influence d'un idiome étranger, on peut citer à juste titre le huzvârèche. L'araméen l'a pénétré, pour ainsi dire, de toutes parts ; sa phonétique, sa grammaire, son lexique offrent à chaque instant les preuves les moins équivoques de l'influence araméenne. S'il pouvait exister une langue mixte, le huzvârèche serait assurément l'un des exemples les plus éclatants de ce phénomène. Mais il ne saurait y avoir d'hybrides de cette sorte et le huzvârèche est en réalité un idiome éranien, tout comme l'anglais, malgré tout ce qu'il tient du français, est un idiome germanique. A côté des éléments araméens que présente la langue de l'époque des Sassanides, celle du Bundehèche compte un certain nombre d'éléments arabes qui trahissent sa forme plus récente ; ce livre est dû peut-être à quelque Parsi érudit qui connaissait à fond la langue de la traduction des livres sacrés (2).

La grammaire huzvârèche a bien perdu de cette cor-

(1) *Grammatik der huzvâreschsprache*, p. 21. Vienne, 1856.
(2) F. Justi. *Der Bundehesch*, Préface, p. viii. Leipzig, 1868.

rection, de cette fidélité aux formes antiques, qui caractérisaient le zend et le perse ancien. Le genre des noms ne peut plus se distinguer par la désinence, le duel a disparu ; l'accusatif n'a pas plus de terminaison que le nominatif ; le génitif, ou, pour mieux dire, l'idée qui répond à ce qu'exprimait l'ancienne forme du génitif, est rendue par l'entremise d'un élément *i*, reste d'un ancien pronom relatif ; l'idée qui répond à l'ancienne forme du datif, est exprimée au moyen de particules, de véritables prépositions. La conjugaison a tout autant souffert. Quoi qu'il en soit, le fond même de la langue est toujours demeuré éranien. C'est ce qui apparaît bien clairement dans ce fait que le huzvârèche possède des vers composés, formés non seulement d'une racine éranienne et d'une préposition éranienne, mais encore d'une racine éranienne et d'une préposition sémitique, d'une racine sémitique et d'une préposition éranienne, et, ce qui est bien plus curieux, d'une racine sémitique et d'une préposition également sémitique. Or l'on sait que le sémitisme, à la différence de l'indo-européen, n'a point de verbes composés ; il n'a aucune forme, par exemple, qui puisse à côté du mot « prendre » répondre à nos formes « apprendre, comprendre, reprendre, entreprendre » et autres semblables.

Il y a peu d'alphabets qui soient moins pratiques que l'alphabet huzvârèche. Un seul et même caractère y exprime souvent plusieurs sons différents et il s'y rencontre un assez grand nombre de ligatures, c'est-à-dire d'agglomérations de plusieurs caractères groupés en un seul signe. Aussi, dans les écrits de linguistique, ne cite-t-on que très rarement les mots du huzvârèche sous la forme de leur propre écriture ; on les transcrit plutôt en caractères latins ou bien encore en caractères hébraïques ou arabes.

V. *Parsi.*

On a donné quelquefois au parsi le nom de « pâzend ». C'était là un nom inexact. Pour les Orientaux modernes zend et pâzend sont des noms de livres, non point de langues, et leur opinion sur ce point paraît parfaitement fondée. A la vérité, le nom de « zend » a prévalu sur tout autre lorsqu'il s'est agi de désigner la langue de l'Avesta, mais celui de « pâzend » n'est point tellement accepté qu'on ne puisse lui substituer le terme bien plus juste de *parsi* « langue des Parses ».

Le parsi fut très certainement contemporain du huzvârèche, mais il lui survécut de plusieurs centaines d'années et fut à la fois langue vulgaire et langue littéraire. Il était parlé d'ailleurs dans une région plus orientale de l'Eran et l'on n'y trouve point cette foule d'éléments araméens que possède le huzvârèche.

La grammaire du parsi est, comme la grammaire du huzvârèche, bien éloignée de cette apparence antique qui distingue le zend et le vieux perse. Il n'est pas éloigné du huzvârèche, mais il se rapproche beaucoup du persan. Toutefois, il l'emporte d'une façon très visible sur ce dernier idiome ; c'est ainsi, par exemple, qu'il conserve beaucoup mieux les éléments pronominaux anciens, et qu'il possède un grand nombre de verbes que le persan ne connaît plus. Eugène Burnouf et F. Spiegel pensent que le parsi a pû être parlé jusqu'à l'époque du poète persan Firdousi, c'est-à-dire jusqu'aux premières années du onzième siècle.

Le parsi n'a point de système particulier d'écriture ; tantôt il emploie les caractères zends, tantôt il se sert des caractères arabes.

Aujourd'hui les Parsis sont établis principalement à

Bombay, à Surat, à Baroda et dans le Gouzérat, au nombre de cinquante mille pour certains auteurs, de quatre-vingt mille pour d'autres, pour d'autres encore au nombre de cent cinquante mille.

VI. *Persan.*

De toutes les langues éraniennes modernes le *persan*, ou « néo-perse », est la plus répandue et la mieux connue. Le persan est un dialecte éranien qui, depuis l'an 1000 environ, devint une langue littéraire. Nous n'avons pas à faire ici l'histoire de sa littérature ; disons seulement qu'elle fut d'une importance considérable. Elle aborda simultanément la poésie, l'histoire, les sciences. Le « Livre des Rois » de Firdousi (fin du dixième siècle et commencement du onzième) est une époque nationale qui peut rivaliser avec les œuvres capitales de bien d'autres langues littéraires (1).

Le persan a adopté les caractères arabes augmentés de quatre lettres « p, tch, j, g (dur) ».

La déclinaison n'existe plus en persan : on joint simplement au nom certaines prépositions lorsqu'il s'agit d'exprimer l'idée du datif et celle de l'accusatif. S'agit-il de rendre l'idée du génitif, on intercale, comme en huzvârèche et en parsi, l'élément *i*, reste d'un ancien pronom relatif : *dast i pusar* « la main de l'enfant », *pusar i man* « mon enfant », ce qui revient à dire : la main qui (est celle de) l'enfant ; l'enfant qui (est) le mien. C'est un procédé syntaxique.

La conjugaison est également simplifiée. Les suffixes personnels sont assez bien conservés — *m* à la première personne du singulier et du pluriel, *d*, pour un « t » plus ancien, à la troisième personne, — mais les temps sont

(1) Mohl. *Firdousi. Le livre des Rois publié en persan avec une traduction française en regard.* Paris, 1838, ss.

traités, comme l'ont été les cas, au moyen de procédés modernes ; le persan, en un mot, est devenu une langue analytique.

Quant au lexique persan, il contient un grand nombre de mots empruntés à l'arabe. Au surplus, à côté de la langue persane littéraire il existe un certain nombre de dialectes populaires (par exemple le mazadaran), qui ont chacun quelques particularités, aussi bien sous le rapport du vocabulaire que sous le rapport de la phonétique et, parfois, sous celui des formes elles-mêmes.

VII. *Ossète, kourde, béloutche, afghan, etc.*

Bien que réunis sous une seule et même rubrique, ces différents idiomes ne sont pas plus rapprochés entre eux dans la famille éranienne que ne le sont tels ou tels des autres idiomes dont nous avons parlé ci-dessus.

La déclinaison de l'*ossète* est plus riche que celle du persan ; sa conjugaison, par contre, est assez analogue à celle de cette dernière langue, et, en somme, il se rapproche davantage des formes éraniennes plus anciennes que l'on rencontre en arménien, en huzvârèche, en parsi. L'ossète est parlé au nord et au sud du Caucase, aux alentours du col de Dariel, c'est-à-dire vers le centre même de cette chaîne de montagnes, et il se divise en un certain nombre de dialectes.

On peut dire d'une façon sommaire que le *kourde* est allié de près au persan, plutôt peut-être aux dialectes de cette langue qu'à l'idiome littéraire lui-même. Sa phonétique semble plus altérée que celle du persan. On compte chez les Kourdes un certain nombre de dialectes dont le principal est le « kourmandji », dialecte occidental, parlé depuis Mossoul jusqu'à l'Asie Mineure.

Le « zaza » est moins pur sous certains rapports, plus pur sous certains autres.

Le *béloutche* se rapproche du kourde. Il contient un assez grand nombre d'éléments étrangers, notamment de mots empruntés à l'arabe.

Quelques auteurs seraient portés à ne pas considérer l'*afghan* — « pachto » ou « pouchtou » — comme une simple langue éranienne ; il faudrait le regarder comme un idiome indépendant, formant une classe par lui-même et apparenté aux langues hindoues tout aussi bien qu'aux langues éraniennes. L'opinion de Fréd. Müller est toute différente ; l'afghan est à ses yeux un dialecte de l'éranien oriental, le descendant d'un ancien dialecte de la Baktriane. Sa conjugaison est inférieure encore à celle du persan, celui-ci, en effet, a conservé certaines formes anciennes du temps présent, tandis que l'afghan les a toutes perdues ; la racine du verbe et le thème du temps présent sont ordinairement identiques chez lui. Quant à son lexique, il contient un certain nombre de mots persans et de mots arabes.

Nous sommes loin d'avoir cité tous les idiomes éraniens modernes. A côté de ceux dont nous venons de parler, et que l'on peut regarder comme les plus importants et les mieux connus, il en est un certain nombre d'autres : la langue des Loures (Bachtiaris et Féilis), dont on ne possède que fort peu de documents, est assez rapprochée du kourde ; la langue des Tâts, au sud-est du Caucase, n'est pas éloignée du persan.

Il est certain, d'ailleurs, que bien des dialectes éraniens autres que le zend, le perse, le huzvârèche, le parsi, ont péri dans le cours des âges. Il se peut que parmi les populations auxquelles les auteurs anciens, notamment les auteurs grecs, ont donné le nom de « scythiques », il s'en trouvât dont la langue fût incontestablement éranienne ; il existe à cet égard quelques

présomptions assez vraisemblables, mais les documents dont on dispose sont encore trop restreints pour qu'il soit possible d'émettre à ce sujet une opinion définitive.

Certaines langues de l'Asie Mineure ont été rangées également au nombre des langues éraniennes, le phrygien, par exemple, que l'on a rapproché plus particulièrement de l'arménien, le lycien, le carien, d'autres encore ; mais cette classification est peut-être prématurée. Il y faudrait plus de réserve. Nous pensons ne devoir dire quelques mots de ces idiomes qu'en énumérant un certain nombre de langues dont le caractère indo-européen est évident, mais dont la position véritable, dans l'ensemble de cette famille linguistique, n'est pas encore établie.

§ 3. Branche hellénique.

De toutes les langues indo-européennes parlées en Europe, le grec est incontestablement celle qui se rapproche le plus du sanskrit et des langues éraniennes. Il est possible, il est même probable, que la connaissance des idiomes indo-européens de l'Asie-Mineure, le phrygien, le lycien et autres, montrera un jour que ces liens sont plus intimes qu'on ne le suppose. Nous reviendrons un peu plus loin sur cette question des degrés de parenté des différentes langues indo-européennes. Ici il nous suffit de prémunir le lecteur contre cette opinion, jadis très répandue et aujourd'hui encore assez courante, que le grec et les langues italiennes forment une branche particulière de la grande famille linguistique à laquelle ils appartiennent. Le grec, sans doute, a des rapports bien intimes avec le latin, mais il a des rapports tout aussi intimes avec les langues indo-européennes de l'Asie, le sanskrit et le perse ; le latin, d'autre

part, est en bien des points plus rapproché des langues celtiques qu'il ne l'est du grec.

Le grec a beaucoup mieux conservé le système des voyelles de l'indo-européen commun qu'il n'a conservé le système des consonnes ; sous ce rapport, il se rapproche tout particulièrement du zend et du vieux perse. Il garde, par exemple, les anciennes diphthongues que le latin, comme le sanskrit, réduit en une voyelle longue. En ce qui touche les consonnes il est moins correct. Un de ses changements les plus caractéristiques est celui des explosives aspirées, *gh*, *dh*, *bh*, en aspirées *kh*, *th*, *ph* χ, θ, φ. On ne saurait dire sous quelle influence se fit cette variation ; le fait est qu'elle est constante. Tandis que le sanskrit dit *dirghas* « long », *bhârami* « je porte », le grec dit δολιχός, φέρω. Loin de maintenir, comme l'a fait le latin, tous les « k » primitifs, il les change souvent soit en *p*, soit en *t*. Tandis, par exemple, que le latin dit *quis*, *quinque*, il dit τίς, πέμπε et πέντε.

Mais c'est lorsqu'il s'agit des « s, y, v » primitifs qu'il s'écarte le plus du type commun et qu'il se montre inférieur à tous les autres idiomes indo-européens, à tous sans exception. Un mot commence-t-il par *s*, il change d'habitude cette sifflante en un esprit rude, que l'on transcrit ordinairement par *h* ; c'est ainsi que ἡδύς « doux » correspond au sanskrit « svâdus », ἑπτά « sept » au latin « septem », ἑκυρός « beau-père » au latin « socer ». Parfois cette sifflante disparaît tout à fait, notamment lorsqu'elle se trouve entre deux voyelles. Le « y » primitif tombe fort souvent, lui aussi, surtout entre deux voyelles. Au commencement des mots il se change en ζ (prononcez « dz ») ou en esprit rude ; c'est ainsi que ζυγόν « joug », ἅγιος « saint » correspond au sanskrit « yugam, yajyas ». Le « v » primitif tombe, en principe, dans le grec classique, ou bien se change en υ. Certains dialectes, ainsi que nous le verrons tout à l'heure, ont

maintenu l'ancien « v » sous la forme du digamma, et disent *nevos*, *voikos*, *ovis* ; mais le digamma ne prévalut point dans l'idiome d'Athènes que les événements politiques rendirent prépondérant et classique.

Moins compliquées que les lois phonétiques du sanskrit, celles du grec sont pourtant assez importantes. Elles reposent, en général, sur une vive tendance à l'assimilation des consonnes d'ordre différent qui viennent à se rencontrer. Le « zétacisme » joue également un grand rôle dans tous les dialectes grecs : les groupes organiques « g+y », « d+y » se changent souvent en ζ. Ainsi Ζεύς correspond à la forme sanskrite « dyâus ». A la fin des mots, le grec n'admet point d'autres consonnes que *s* ou *n* ; ainsi tous les « m » qui déterminent l'accusatif singulier se changent en ν ou bien tombent : le grec dit φέροντα « portant », ναῦν « vaisseau », tandis que le sanskrit, plus correct, dit « nâvam ».

La déclinaison du grec est bien conservée. Elle a perdu, à la vérité, l'ablatif singulier, mais elle a gardé l'ancien locatif, tant au pluriel qu'au singulier. Parfois, celui-ci sert de datif ; μητρί « à la mère », νέκυι « au mort », ποιμένι « au berger », mais la forme n'a rien de commun avec celle du datif, et elle n'en a pris la signification que par la suite des temps. Le locatif du pluriel est en *si* : ναυσι « dans les vaisseaux », Ἀθήνησι, Ὀλυμπίασι. Les grammaires classiques en font autant de datifs, ce qui est tout à fait inexact. Le grec possède aussi, sous la forme unique φι la désinence de l'instrumental singulier « bhi » et de l'instrumental pluriel « bhis », que beaucoup d'autres langues indo-européennes ont perdue. La grammaire ancienne traitait cette désinence φι de simple addition ; en réalité, il s'agit d'un véritable cas. Quant au duel, le grec le possède en partie ; il en a perdu le locatif et le véritable génitif.

Si nous passons à la conjugaison, nous voyons que le grec possède l'ancienne voix intransitive (λύομαι, λύεται) que les langues italiques, celtiques, slaves et lettiques ont perdue ; c'est un grand avantage. D'autre part, il conserve assez bien les six temps anciens. Il s'est créé, d'ailleurs, un certain nombre de temps nouveaux, entre autres, un plus-que-parfait tiré du parfait redoublé. En somme, on peut dire que le grec ancien est assez fidèle au type commun d'où il est sorti, lorsqu'il s'agit de la formation des mots, mais qu'en bien des points de sa phonétique il s'en est singulièrement écarté.

Les différences dialectiques sont presque toutes d'ordre phonétique. On compte un grand nombre de dialectes grecs, mais ils se groupent tous assez facilement sous quatre formes spéciales, l'éolien, le dorien, l'ionien, l'attique. Parfois même on ne fait de ces quatre dialectes que deux grandes classes : l'éolien et le dorien formeraient un groupe à eux deux, l'ionien et l'attique en formeraient un autre.

L'éolien proprement dit (1) était parlé en Asie Mineure : Alcée, Sapho écrivirent dans le dialecte de Lesbos. Il possède le digamma correspondant au « v » antique et redouble volontiers les consonnes liquides ; il dit, par exemple, ἐμμί « je suis ». Là où l'ionien dit η pour « â » plus ancien, il conserve souvent cette dernière voyelle. Entre autres faits également caractéristiques, il possède un plus grand nombre de verbes en μι que la langue ordinaire : il dit, par exemple, φίλημι « j'aime ». Le béotien appartenait au même groupe. Il a laissé peu de productions littéraires; il use du digamma, condense les diphthongues en une simple voyelle longue, dit également « â » là où l'ionien dit η et met souvent

(1) Ahrens. *De græcæ linguæ dialectis*. 2 vol. Gœttingen, 1839-1843.

un δ à la place du ξ ordinaire : l'attique dit Ζεύς, ζυγόν, le béotien dit Δεύς, δυγόν. Le thessalien, dont il n'y a que fort peu de restes, faisait partie lui aussi du groupe éolien ; on le tenait, à Athènes, pour un dialecte assez barbare.

Le dorien était parlé dans presque tout le Péloponnèse, en Crète et dans les colonies grecques de Sicile, de Libye, de l'Italie méridionale. Pindare écrivit en dorien. On y distingue deux sous-dialectes, l'un plus rigoureux que l'autre. Le dorien possédait le digamma et conservait le « t » ancien que la langue classique changeait en σ ; il disait, par exemple, δίδωτι « il donne », ϝίκατι, ϝείκατι « vingt ».

L'ionien eut deux périodes, une période ancienne et épique, celle de la langue d'Homère et d'Hésiode, une période plus récente, celle d'Hérodote. On le parlait en certaines localités de l'Asie Mineure, en Attique et dans un grand nombre d'îles.

Bien des auteurs rattachent l'attique à l'ionien ; il en diffère fort peu, en effet, et peut être envisagé comme un dialecte ionien. C'était la langue d'Athènes, la langue d'Eschyle, de Sophocle, d'Euripide, d'Aristophane, de Thucydide, de Démosthène ; c'est elle qui prédomina, par la suite des temps, sur les autres dialectes, c'est elle qu'ont toujours en vue les réformateurs de la langue grecque moderne.

Chaque dialecte, ainsi que nous l'avons dit, eut sa littérature ; toutefois, grâce aux circonstances politiques, le dialecte attique gagna peu à peu une prédominance incontestée et devint la langue écrite commune, le dialecte commun ἡ κοινὴ διάλεκτος. Cette extension un peu factice fut précisément la cause de sa décadence et de sa corruption. Le *dialecte commun* ne fut plus dans la bouche des Grecs étrangers à l'Attique, et encore moins dans la bouche des barbares qui l'acceptèrent, ce qu'il

était à Athènes ; il devint petit à petit le *byzantin*, la langue byzantine du moyen âge.

Le grec du moyen âge va de l'an mil à l'an seize cent, environ. J. Psichari a démontré que les textes des auteurs de cette époque donnent une image fidèle de la langue, qui n'était en aucune façon un mélange de formes contemporaines et de formes plus anciennes, comme on l'a dit. Plus on remonte, plus les formes modernes sont rares ; plus on approche du dix-septième siècle, plus elles sont fréquentes (*Essais de grammaire historique néo-grecque*, 1886).

Au grec du moyen âge succéda le grec moderne ; on lui donna le nom de *romaïque*, souvenir du vieil empire romain ; c'est un nom malheureux, qui peut prêter à confusion et dont nous ne nous servirons pas.

Dans le grec moderne, il ne faut pas chercher d'anciennes formes dialectales ; il provient simplement de l'ancienne langue écrite commune, langue ayant une phonétique et une morphologie uniformes.

Le grec actuel comprend un grand nombre de dialectes qui diffèrent assez sensiblement les uns des autres, et que l'on rencontre non seulement dans les îles, mais encore sur le continent, comme par exemple le tsaconien parlé en pleine Morée. Quant à la langue littéraire, la langue commune, elle est réellement peu éloignée de la langue classique d'il y a deux mille ans.

C'est précisément cette grande ressemblance qui a donné à quelques Hellènes l'idée d'une rénovation de leur langue, fondée sur le retour aux formes et aux expressions mêmes de la langue de Thucydide. Aucune tentative ne pouvait être moins pratique ; celle-ci rentre dans le domaine de la fantaisie et de la curiosité. Le grec moderne diffère bien peu, sans doute, du grec ancien, mais il en diffère très nettement. Il a perdu les formes du duel et le datif : ce dernier cas n'est employé

que dans la littérature la plus élevée et il serait pédant de s'en servir dans la conversation, ou même dans la littérature courante. L'ancien infinitif en ειν n'est plus employé, non plus, que dans la littérature pseudo-classique ; on le remplace ordinairement par une forme conjonctive : « je veux venir » s'exprime par cette formule analytique θέλω να ἔλθω « je veux que je vienne ». Le futur est également analytique : il est formé, entre autres façons, du présent précédé d'une conjonction. Il y a bien d'autres exemples d'analytisme dans la conjugaison du grec moderne, mais nous en avons dit assez pour ne pas insister davantage (1).

Le grec moderne se distingue encore du grec ancien par un fait qui, pour ne pas se rapporter aux formes elles-mêmes des mots, n'en est pas moins très important : l'accent, chez lui, tient lieu de prosodie. En d'autres termes, la syllabe longue du grec moderne est celle qui porte l'accent, la syllabe brève est celle qui n'est pas accentuée. Ce phénomène n'est point particulier à la langue grecque ; nous verrons dans le chapitre consacré aux langues germaniques qu'il constitue l'une des caractéristiques de l'allemand moderne : en moyen haut-allemand, aux douzième, treizième, quatorzième, quinzième siècles, la syllabe radicale était tantôt longue, tantôt brève, tandis que dans l'allemand moderne, portant l'accent, elle est toujours longue. C'est là un phénomène tout moderne.

Ce n'est pas seulement en Grèce que la langue grecque est parlée ; en Turquie, elle occupe la Thessalie — confinant vers l'ouest à l'albanais, vers le nord au bulgare ; — elle longe toute la côte septentrionale de la mer Egée et tout le contour de la mer de Marmara, aussi bien en Europe qu'en Asie. Elle atteint parfois dans

(1) Egger. *De l'état actuel de la langue grecque et des réformes qu'elle subit.* Paris, 1868.

l'intérieur même des pays une certaine profondeur : elle s'étend, par exemple, jusqu'à Andrinople. Dans l'île de Candie, elle domine presque totalement ; on n'y rencontre plus qu'une seule enclave de langue turque qui soit importante : c'est vers le centre de l'île. On évalue à plus d'un million le nombre des Grecs de Turquie. En Russie, il existe sur la côte nord de la mer d'Azov, entre la Tauride et les Cosaques du Don, deux enclaves de langue grecque. Elle occupe également les trois côtes de l'Asie Mineure depuis la région opposée à l'île de Chypre jusqu'à l'embouchure du Kysyl Irmak dans la mer Noire.

Nous avons à parler maintenant d'une question un peu accessoire, mais qui a cependant son intérêt, celle de la prononciation du grec moderne et du grec ancien.

Six caractères, dont trois sont simples et trois sont doubles, répondent dans le grec moderne à notre voyelle « i ». Ce sont *ê*, *i*, *u* (η, ι, υ) et *ei*, *oi*, *ui*. Les autres voyelles se prononcent comme en français, *o*, *ô*, etc. D'autre part, les groupes αυ, ευ, ηυ, ου se prononcent « av, ev, iv, ov ».

En ce qui concerne les consonnes nous voyons que *th* (θ) correspond au « th » des Anglais (de « three ») ; *d* (δ) au « th » doux (de « thus, the »). Le *ph* (φ) vaut « f », et *kh* (χ) vaut le « ch » allemand de « noch, nach, buch » ou celui de « ich, fechten » selon les voyelles qui l'avoisinent. Le *g* (γ) se prononce comme le « y » français s'il précède un « e » ou un « i ». Il se trouve, on le voit, une grande différence entre cette prononciation et la prononciation dite classique, attribuée à Erasme.

Il existe toute une école d'hellénistes aux yeux desquels la prononciation du grec moderne devait être appliquée au grec ancien et qui mènent en faveur de cette théorie une propagande bruyante. Elle n'est rien moins que scientifique. Prononcer le grec ancien à la façon mo-

derne, c'est une erreur, dit très justement Schleicher, due à une complète méconnaissance des lois de la vie linguistique et de la phonétique. *A priori*, en effet, cette théorie n'est point soutenable : en fait, elle ne supporte pas la critique.

La comparaison même du grec ancien et des autres idiomes indo-européens nous enseigne que les sons *ê*, *i*, *u* (η, ι, υ) correspondaient aux voyelles « â, i, u (prononcez ou) » et qu'elles étaient, dès lors, parfaitement distinctes ; ce n'est que par la suite des temps qu'elles se sont toutes les trois confondues en un seul et même son, « i ». La transcription d'η grecs par des *e* latins et d'*e* latins par des η grecs démontre très clairement que le son de l'ancien η grec n'était pas celui d'un « i » ; on trouve, par exemple, κηνσωρ, Αυρηλιους, pour « censor, Aurelius ». Il n'est pas douteux davantage que la voyelle υ n'ait eu autrefois en grec le son de notre « ou », de la voyelle latine « u » ; à l'époque classique il avait la valeur de « u » français, et la diphthongue ου, c'est-à-dire « o+u », se réduisit à cette même époque en une simple voyelle υ, notre « ou » français, la voyelle latine « u ». C'est ainsi que le grec transcrit Τιτιους, τουομ, κιρκουιτουμ, les mots latins « Titius, tuum, circuitum ». Il est tout aussi certain que primitivement le *b* grec (β) se prononçait comme notre « b » et non pas « v » comme il se prononce aujourd'hui. Dans les auteurs grecs le bêlement des moutons est rendu par βη, βη, que l'on ne pourrait prononcer sans ridicule « vi, vi ». A une certaine époque, les Grecs ont bien transcrit le « v » latin par leur β mais auparavant ils l'avaient rendu par ου : on trouve par exemple, Ουαρων, Ουαλεριος, Ουεργιλιος, pour « Valerius, Vergilius, » etc. ; le changement de « b » en « v » fut peut-être assez précoce, en quelques dialectes du moins, mais partout le « b » avait eu jadis sa propre et véritable prononciation. A l'époque où les

Grecs commencèrent à transcrire des noms latins, leur ε était loin d'avoir toujours et en tous lieux le son qu'il a actuellement chez eux. A cette époque on l'emploie régulièrement lorsqu'il s'agit de transcrire un « b » latin, et ce n'est qu'en concurrence avec « ou » et « o » qu'on l'emploie, à la même époque, pour rendre le « v » latin (1).

Enfin il ne saurait y avoir de doute sur la prononciation ancienne des aspirées *ph*, *th*, *kh* (φ. θ. χ). Elles avaient bien la valeur de « p, t, k » aspirés, c'est-à-dire « p+h, t+h, k+h. Le « th » dur anglais, le « f », les deux « ch » allemands n'ont rien à faire ici. Aujourd'hui ces consonnes φ. θ. χ sont bien des sifflantes, mais c'étaient autrefois des aspirées. On peut le démontrer de diverses façons (2). La mobilité de l'aspiration qui accompagne la simple explosive « p, t, k » (π. τ. κ) est une première preuve : le redoublement de la racine θε donne τίθεμεν « nous plaçons » ; de même le redoublement de *ph* (φ) se fait par un simple *p*, celui de *kh* (χ) par un simple *k*. C'est ainsi que le sanskrit redouble « dh, bh » par les simples consonnes non aspirées « d, b » et dit « dadhâmi, je pose ; babhâu, j'ai brillé ». Dans les formes telles que τρέφω « je nourris » et θρέψω (= θρεπσω) « je nourrirai », nous voyons avec la même clarté combien l'aspiration est mobile ; ici, comme dans le cas précédent, il est évident que le *ph* et le *th* ne sont pas des sifflantes, mais bien de pures consonnes aspirées. Le sanskrit dit d'une façon tout analogue « badhnâmi, je lie ; bhatsyâmi, je lierai ». On peut encore remarquer que certains dialectes déplacent aisément l'aspiration dans le corps d'un mot : le grec commun dit ἐνταῦθα, χιτών. dialectalement on dit ἐνθαῦτα, πιθών. Dans la bouche des

(1) G. Curtius, *Grundzüge der griechischen Etymologie*. Quatr. édit., p. 571. Leipzig, 1873.
(2) *Ibid.*, p. 416.

barbares qu'il met sur la scène, Aristophane remplace par de simples consonnes non aspirées « p, t, k » les aspirées grecques « ph, th, kh » ; c'est encore là un argument décisif. Un autre argument du même genre se puise dans la façon dont la vieille langue populaire latine rend ces mêmes aspirées du grec : elle les dépouille simplement de leur aspiration. Au quatrième siècle même, le gothique rend par un « k » les « kh » (χ) du grec. Enfin, bien des dialectes du grec moderne ont une simple et pure consonne non aspirée à la place de la consonne aspirée de la langue littéraire ; il est hors de doute que ces dialectes reflètent ici une époque très ancienne, ce qui, d'ailleurs, n'est point un cas fort rare. En somme, il est incontestable que les anciennes aspirées grecques avaient la valeur de « p+h, t+h, k+h » et qu'elles n'ont passé à l'état de sifflantes que dans la suite des temps.

Au surplus, ce serait une entreprise vaine que de chercher à déterminer l'époque du changement de prononciation de la langue grecque. Nous parlons, bien entendu, d'une façon générale. Deux facteurs ont contribué à ce changement, ou, pour mieux dire, à ces changements divers : le temps et le lieu ; telle variation s'opérait ici à telle époque, qui ne devait s'opérer ailleurs que beaucoup plus tard et qui, ailleurs encore, était déjà depuis longtemps un fait accompli. Il ne peut donc être question, dans une étude sur l'ancienne prononciation du grec, que de monographies toutes particulières. Par la suite on pourra rassembler ces recherches spéciales et voir s'il est possible ou non d'en tirer un enseignement plus général ; mais jusqu'à ce moment il faudra s'en tenir à la prononciation dite érasmienne, bien qu'elle soit souvent fautive, et rejeter la prononciation moderne, bien plus fautive encore.

§ 4. Branche italique.

Jusqu'au jour où les bases de la grammaire comparée des langues indo-européennes furent définitivement établies, on put croire que le latin et les autres idiomes de l'Italie ancienne qui lui étaient alliés, avaient leur source dans la langue grecque. On citait même celui des dialectes grecs qui avait donné naissance aux anciennes langues italiques, l'éolien. Ce fut précisément un des résultats de la grande œuvre de Bopp, de montrer que le latin ne procédait pas plus du grec que celui-ci ne procédait du sanskrit et que tous les trois descendaient d'une mère commune, de la langue qui avait également donné naissance aux idiomes éraniens, slaves, lettiques, germaniques et celtiques.

La grammaire comparée nous apprend, en effet, que le latin possède une foule de formes plus anciennes que les formes qui leur correspondent dans la langue grecque. S'agit-il de la phonétique, nous voyons, par exemple, que le latin conserve au commencement des mots les *s* antiques que le grec change en esprits rudes : il dit *septem*, *sex*, *socer*, alors que le grec dit « hepta, hex, hekuros » (ἑπτά, ἕξ, ἑκυρός) ; — Nous voyons qu'il conserve l'ancienne demi-voyelle « y » (qu'il représente par le signe *j*) là où le grec la change en ζ (prononcez « dz »), ou en esprit rude : *jecur jugum* sont plus purs que ἧπαρ, ζυγόν ; — nous voyons qu'il conserve les « k » primitifs que le grec change souvent en « p », en « t » ; *quinque*, *quis* l'emportent sur πέντε, τίς. Naturellement c'est par la confrontation méthodique avec les autres idiomes indo-européens que l'on peut démontrer la correction de ces différentes formes grecques qui leur correspondent. S'agit-il de la déclinaison, nous voyons que le latin a conservé l'ablatif singulier que le grec ne connaît plus. S'agit-il de la conjugaison, nous voyons

que le suffixe de la seconde personne du pluriel est d'une forme plus correcte en latin qu'en grec : *estis* « vous êtes » se rapproche plus de la forme organique « astasi » que ne le fait le grec ἐστέ, le lithuanien « este », le sanskrit « stha ». Par contre, le grec l'emporte souvent sur le latin : dans la conjugaison, notamment, où il a mieux conservé les temps anciens. L'un et l'autre ont aussi leurs côtés forts et leurs côtés faibles, et aucun d'eux, en fin de compte, ne peut se dire plus correct, plus pur, plus antique que son congénère.

Nous allons traiter successivement des anciennes langues italiques, latin, ombrien, etc., et des langues romanes, ou novo-latines, actuellement parlées dans l'Europe sud-occidentale et sur le cours du bas Danube.

I. *Anciennes langues italiques.*

Le *latin* est le grand représentant de cette branche. A côté de lui, l'*osque* et l'*ombrien* jouent un rôle peu considérable, mais on ne saurait les négliger entièrement. Il se parlait d'ailleurs en Italie un certain nombre d'autres langues appartenant à cette même famille. Nous les passerons sous silence, parce que l'on n'en sait que fort peu de chose.

Nous ne parlerons de l'étrusque qu'après avoir terminé notre examen des différents rameaux de l'indo-européen.

Les formes du *vieux latin*, que l'on rencontre jusqu'au milieu du troisième siècle avant notre ère, c'est-à-dire avant l'époque de la première guerre contre les Carthaginois, et qui nous sont connues par un certain nombre d'inscriptions, ne diffèrent pas essentiellement des formes latines dites classiques. Les différences qui se rencontrent dans la langue des deux périodes concernent surtout la phonétique et en particulier les voyelles.

On peut dire que le latin classique se distingue tout d'abord du vieux latin par un besoin très marqué de réduire en voyelles simples les anciennes diphthongues ; c'est plus qu'une tendance, c'est un caractère très prononcé. La diphthongue *au* est la seule qui se soit à peu près maintenue, les autres se sont presque toujours condensées en une voyelle longue. Tandis, par exemple, que le vieux latin disait *loumen*, *jous*, *oinus*, *oitile*, *ploeres*, *ceivis*, *leiber*, *veicus*, le latin classique disait *lumen*, *jus*, *unus*, *utile*, *plures*, *civis*, *liber*, *vicus*. L'ancienne diphthongue *ai* est définitivement devenue *ae* au temps des Gracques, cent trente ans avant notre ère, et cet *ae* se change à son tour en *e* long, d'abord dans la langue populaire, avant notre ère déjà, puis dans la langue cultivée, trois ou quatre cents ans plus tard (1).

Certains changements de voyelles simples qui s'effectuent dans le passage de l'ancienne période latine à la période classique, sont d'une importance moindre, mais il faut encore les regarder comme caractéristiques. C'est ainsi, par exemple, que la voyelle *o* devient parfois *e* : les anciennes formes *vorto*, *voster*, deviennent *verto*, *vester* ; *u* devient *i* : *optumus*, *decumus*, *mancupium* se changent en *optimus*, *decimus*, *mancipium* : *i* devient *e* : *navim* se change en *navem*.

Ces différentes mutations, ainsi qu'un assez grand nombre de variations analogues, ne sont point régies, sans doute, par des lois spéciales, elles ne sont pas aussi régulières que l'est la condensation des anciennes diphthongues en une voyelle simple ; mais elles affectent cependant une certaine allure bien caractérisée à laquelle ne peuvent se tromper ceux qui sont quelque peu familiarisés avec les formes classiques ordinaires.

(1) Corssen, *Ueber Aussprache, Vokalismus und Betonung der lateinischen Sprache*, Deuxième édit., t. I, p. 695. Leipzig, 1868.

Les principes d'euphonie particuliers aux voyelles latines sont peu nombreux.

Au demeurant, le tableau est des plus simples, et il se rapproche beaucoup du système des voyelles grecques qui n'a fait simplement que mieux conserver les anciennes diphtongues.

A l'égard des consonnes, le latin est bien plus fidèle que le grec au système primitif. Le lithuanien seul, de tous les idiomes indo-européens, garde plus fidèlement que le latin l'unique sifflante *s* de l'indo-européen commun. Le latin, en effet, la change parfois en *r* entre deux voyelles, comme dans *generis*, génitif de *genus* (dont le thème est « genes » comme l'indiquent le sanskrit et le grec), ou encore à la fin des mots, comme dans *arbor*, dont la plus ancienne forme est *arbos*. Mais cette seule et unique variation est bien moins importante que ne le sont toutes ces créations de sifflantes nouvelles particulières au grec, aux langues slaves, aux langues éraniennes, au sanskrit.

Tandis que le grec changeait en aspirées fortes « ph, th, kh », (φ, θ, χ), les aspirées faibles de la langue commune indo-européenne « bh, dh, gh », le latin, au milieu des mots surtout, les rendait, en principe, par l'explosive correspondante non aspirée *b*, *d*, *g*. Il dit, par exemple, *nubes* ,*lingo*, tandis que le grec dit νεφος, λείχω. Cependant il traita encore ces aspirées primitives de deux autres façons, surtout au commencement des mots : il les changea soit en *h*, soit en *f*. C'est ainsi que *fero* « je porte » correspond au grec φέρω, au sanskrit « bharâmi ». Parfois les deux formes coexistent : *hordeum* et *fordeum* « orge » *horda* et *forda* « vache pleine ». On a expliqué de différentes façons ce changement des aspirées primitives en *h* et en *f*, mais la question n'est pas encore élucidée et nous devons nous contenter de mentionner le fait pur et simple.

Une autre particularité de la phonétique latine est le changement en *l* d'un ancien *d* : *lacrima* « larme », *levir* « beau-frère », *lingua* « langue », *olere* « sentir » ont tous remplacé par *l* un *d* plus ancien. C'est ainsi que s'expliquent un certain nombre de formes doubles : *impelimenta* et *impedimenta*, *delicare* et *dedicare*, *olere* à côté de *odor*.

Les consonnes latines se prêtent d'une façon assez délicate aux lois de l'assimilation, au moins d'une assimilation élémentaire. Souvent l'assimilation n'est que partielle : *actus*, par exemple, a un *c* pour un *g* comme l'indique *ago*, mais parfois elle est complète ; c'est ainsi que dans *summus* le groupe « mm » est pour « pm » comme l'indiquent *super*, *supremus*. Un mot vient-il à commencer par un groupe de deux consonnes, la première de ces consonnes disparaît souvent : *notus*, *nomen* étaient autrefois précédés d'un « g » comme en témoignent les composés « cognosco, cognomen, ignotus ». Au commencement des mots, également, le groupe *dv* peut se changer en *b* : *bis* et *bonus* sont pour des formes plus anciennes *dvis* et *dvonus* ; *bellum* et *dvellum* coexistent.

Nous dirons ici quelques mots de la prononciation latine. C'est une question que bien des personnes ont essayé de résoudre, mais sans succès, faute de méthode. Aujourd'hui, l'on peut dire d'une façon générale qu'elle est tranchée ; l'ouvrage de Corssen, que nous citions un peu plus haut, a rassemblé tous les résultats acquis à ce jour et que l'on peut, sans témérité, regarder comme concluants. On s'accorde sans peine pour la prononciation d'une bonne partie des consonnes latines « p, b, f, d, m, n, r, l », etc., de celles-ci nous ne parlerons pas. C'est seulement sur les points qui peuvent encore paraître douteux que nous allons porter notre attention.

On reconnaît généralement que devant les voyelles « a, o, u », et devant les consonnes, le *c* latin avait la

même valeur que le « k » ; mais devant un *e*, devant un *i* comment était-il prononcé ? Avait-il en cette occurence le son de « tch » que lui donnent les Italiens, celui de « ts » que lui donnent les Allemands, celui de « ç » que lui donnent les Français ? Disait-on « Tchitchero, Tsitsero, Çiçero » ? Nous avons plus de renseignements qu'il n'en est besoin pour répondre à cette question, et la transcription en latin des mots étrangers, aussi bien que celle des mots latins en langue étrangère, doit lever tous les doutes. Les Goths, par exemple, empruntant aux Latins les mots *lucerna, carcer, acetum*, en on fait *lukarn, karkara, aikeits ;* les Grecs citaient sous la forme πατρικους. κηνσωρ, κεντυρια les mots latins *patricius, censor*, etc. Par contre, en tous les temps les Latins ont rendu par leur *c* le κ de la langue grecque (comme dans les mots *cerasus, Cimon, Cecrops*), et Corssen conclut avec juste raison que jusqu'au sixième et même au septième siècle de notre ère, le *c* latin avait la valeur de « k » devant toutes les voyelles. Les grammairiens latins ne disent point, d'ailleurs, que la prononciation de cette consonne ait jamais été différente selon qu'elle se trouvait suivie de telle ou telle voyelle, et l'on peut assurer que si le *c* latin fut changé en sifflante devant les « e » et les « i » avant le septième siècle de notre ère, cela n'eut lieu que dans la langue populaire, dans les patois provinciaux.

Devant la voyelle « i », suivie elle-même d'une autre voyelle (justitia, servitium), *t* ne fut également sifflé que très tard, en latin du moins. En osque et en ombrien ce phénomène fut plus précoce, mais ce n'est qu'au cinquième siècle qu'il fut régulièrement reçu dans la bonne prononciation latine, bien qu'on en ait trouvé des traces remontant au troisième siècle de notre ère.

On peut sans crainte, également, prononcer le *g* placé devant les voyelles « i, e » comme on le prononce lorsqu'il est suivi des voyelles « a, o, u ». En aucun cas il ne

faut lui donner la valeur du « j » français. A un moment donné il se changea souvent en *i*, mais ce ne fut que sur le tard et dans la langue populaire.

L'aspirée *h* se fit peut-être entendre à une certaine époque avec quelque force, mais elle perdit peu à peu de sa propre valeur, et nombre de mots latins l'ont tout à fait laissée tomber, *anser*, par exemple, dont la racine est la même que celle du grec χήν « oie ».

La prononciation de j (« jus, jugum ») n'est point douteuse ; elle n'a jamais été celle du « j » français de « jeu, jour », mais bien celle du « y » de « yeux, yeuse ». Le témoignage de Priscien, qui vivait au sixième siècle, ne laisse aucun doute sur ce fait.

La réforme de la prononciation latine serait, somme toute, une chose fort possible ; ajoutons qu'elle serait désirable. Mais se réalisera-t-elle jamais, nous ne pouvons l'espérer. Il est bon, au moins, que l'on connaisse ce qu'était la prononciation aux temps classiques du latin et surtout que l'on ne cherche pas à faire prévaloir aux dépens des autres systèmes l'une des trois prononciations acceptées en France, en Italie, en Allemagne, ou quelque autre système encore ; ils sont tous défectueux au même titre.

D'ailleurs, cette réforme devait être accompagnée de l'observation des lois d'accentuation. Le latin a connu successivement deux procédés d'accentuation. Le premier ne doit point nous arrêter. Le second est celui de l'époque classique. Fondé sur la quantité même des syllabes, on peut dire qu'il est d'une grande simplicité. Son principe est celui-ci : l'avant-dernière syllabe est-elle longue, comme dans *canamus*, c'est sur cette avant-dernière syllabe que tombe l'accent ; cette même syllabe est-elle brève, comme dans *ducere*, l'accent tombe sur la syllabe qui précède l'avant-dernière, c'est-à-dire sur l'antépénultième. Nous supposons ici que le mot a trois syllabes ou

plus de trois syllabes ; s'il n'en a que deux, l'accent tombe sur l'avant-dernière, qu'elle soit longue, comme dans *fecit*, *nobis*, ou brève, comme dans *deus*, *tener*.

L'accent peut donc varier de place dans la déclinaison et dans la conjugaison, selon le nombre de syllabes : dans *amabimur* « nous serons aimés », il tombe sur l'antépénultième, qui est longue ; dans *amabimini* « vous serez aimés » il tombe sur l'antépénultième, qui est brève. Dans ces deux exemples, en effet, l'avant-dernière est brève ; or c'est la quantité de cette syllabe qui décide, avons-nous dit, de la place de l'accent, sans qu'il y ait à s'inquiéter de la quantité des autres syllabes. Cette loi est très importante, car nous verrons que l'accent latin joue un rôle capital dans la formation des langues romanes, notamment de la langue française ; la facture même des mots français est en rapport avec la position de l'accent latin.

Revenons à la grammaire latine que nous avons quelque peu abandonnée pour parler de la prononciation et de l'accentuation.

Des trois nombres, le latin a perdu le duel ; le grec l'a conservé, en partie du moins, et sous ce rapport il est supérieur au latin. En ce qui concerne les cas, le grec et le latin sont tantôt supérieurs, tantôt inférieurs l'un à l'autre. Nous avons dit que le grec avait perdu l'ancien ablatif du singulier : le latin l'a conservé. La désinence organique de ce cas était « t » pour les thèmes qui se terminaient par une voyelle ; en latin ce « t » final se change en *d*. De là les anciennes formes *sententiad*, *preivatod*, *magistradud*, *marid*. Au surplus, ce *d* disparut de bonne heure. La forme organique du datif singulier était « ai » que le sanskrit réduit en « ê » ; de là les anciennes formes *populoi*, *romanoi*, devenues par la suite « populo, romano ». L'ancien locatif avait pour forme organique « i » : le latin ne l'a pas toujours perdu. Il est vrai

qu'il en fait un *î* long, mais cet allongement n'est dû qu'à une cause secondaire dont nous n'avons pas à nous occuper ici. En somme *domi, humi, belli* sont de véritables locatifs (domi agere ætatem, procumbit humi, belli domique), et la grammaire classique les traite sans raison de génitifs. Au pluriel nous avons à remarquer la disparition totale du locatif, qui, ainsi que nous l'avons vu, persiste encore en grec.

Quelques mots sur la conjugaison. En général les désinences qui indiquent la personne sont assez bien conservées en latin. Toutefois de l'ancienne terminaison *mi* « je » du temps présent, il ne reste plus le trace que dans la forme de l'indicatif *sum* « je suis ». Des temps primitifs le latin a conservé le présent, parfois le parfait redoublé (cecinimus « nous avons chanté »), peut-être quelques traces de l'aoriste simple. En somme, cela était fort peu, et il lui a fallu recourir à de nouvelles formations. Le parfait en *si* (luxi, dixi), le parfait en *ui* ou en *vi* (monui, amavi), appartiennent à ces nouvelles formes composées ; il en est de même de l'imparfait en *bam*, du futur en *bo* (amabam, amabo) et d'un certain nombre d'autres formes analogues. C'est là un sujet que nous ne faisons qu'indiquer, en ajoutant que parmi les anciennes langues indo-européennes, le latin est une de celles qui ont donné naissance au plus grand nombre de ces formations nouvelles, dont quelques-unes, sans doute, peuvent paraître superflues.

Il est pourtant une de ces formations, celle du médio-passif, que nous ne pouvons passer sous silence. Dans les langues italiques, comme dans les langue celtiques, on fabriqua une voix moyenne, qui plus tard prit le sens passif, en adjoignant au verbe le pronom réfléchi : *amor* est pour une forme plus ancienne « amos » qui, elle-même, est pour « amo-se ». Le lithuanien, lui aussi, s'est créé un moyen par le même procédé.

De toutes les langues italiques, sœurs du latin, qui devaient dans le cours des temps disparaître peu à peu devant lui, l'osque et l'ombrien sont les plus importantes. L'*ombrien* était parlé au nord-est de la Péninsule, et on admet généralement que le dialecte volsque s'en rapprochait.

L'*osque* était parlé au sud et avait plutôt pour allié le dialecte sabellique. L'ombrien, l'osque, le latin sortaient tous d'une souche commune, et aucun d'eux n'avait précédé les deux autres ; mais la comparaison de leurs formes, de leur phonétique, montre que de ces trois langues l'osque se rapprochait plus particulièrement du type commun qui leur avait donné naissance et que l'ombrien s'en écartait encore plus que le latin.

L'*osque* était parlé dans le Samnium, en Campanie ainsi que dans les pays avoisinants, et il nous est connu par quelques inscriptions assez importantes, les tables de bronze d'Agnone et de Bantia, la pierre d'Abella. L'osque se distingue particulièrement du latin et de l'ombrien par son soin à garder les anciennes diphthongues, et là où le latin remplace par un *i* un « a » plus ancien il conserve cet « a » : il dit, par exemple, *anter*, tandis que le latin dit *inter*. Ces deux caractères d'antiquité ne sont pas les seuls que présente son système vocalique, mais nous pouvons les citer comme très frappants. En ce qui concerne le consonnes, il est parfois inférieur au latin, mais souvent aussi il lui est supérieur. Il se montre inférieur, notamment, en remplaçant par des *p* des « k » primitifs ; il dit, par exemple, *pam* quand le latin dit « quam ». Devant un « t » il remplace les « k » par un *h* : il dit, par exemple, *Ohtavis* tandis que le latin dit « Octavius ». Mais il montre une spériorité réelle en bien des cas. En principe, par exemple, il ne change point les *s* en *r* comme nous l'avons vu faire au latin ; il évite, de même, un certain nombre d'assimilations : il dit *kenstur*,

par exemple, là où le latin dit « censor » pour « censtor ». Une particularité phonétique qui le distingue du latin consiste en ce fait que dans le corps des mots il change souvent les aspirées organiques en *f*, ce que le latin ne fait guère qu'au commencement des mots ; il dit, par exemple, *sifei*, tandis que le latin dit « sibi ».

L'*ombrien* nous est connu par un monument fort important, les tables de bronze dites « tables eugubines » du lieu de leur découverte, Gubbio, l'ancien Eugubium. Trouvées au milieu du quinzième siècle, les tables eugubines ont longtemps exercé la sagacité et la divination des anciens philologues : il était réservé à Th. Aufrecht et Kirchhoff d'amener leur déchiffrement à un résultat vraiment satisfaisant, d'exposer leur grammaire d'une façon scientifique et de publier enfin sur la langue ombrienne un ouvrage dont tous les écrits plus modernes demeurent tributaires (1). Le système des voyelles ombriennes est plus rapproché du système latin que ne l'est celui de l'osque. L'ombrien est porté, plus encore que le latin, à réduire les anciennes diphthongues en une seule voyelle, et, chose plus grave encore, il laisse tomber volontiers bien des voyelles : il dit, par exemple, *nomne* tandis que le latin dit « nomini ». Parfois, comme l'osque, il change les « k » primitifs en *p* ; il dit, par exemple, *pis* quand le latin dit « quis » ; tout comme l'osque il rend par *f* des aspirées primitives que le latin rend par l'explosive simple : le latin dit « tibi, ibi », l'ombrien dit *tefe*, *ife*. Comme l'osque, également, il change le groupe « kt » en *ht* : *rehte* correspond au latin « recte ». En certaines circonstances le « d » primitif se change en un *r* dont la valeur semble particulière et que l'on figure habituellement par un point placé en dessous de ce ca-

(1) *Die umbrischen Sprachdenkmaler*. Berlin, 1849-51. — André Lefèvre. *Les dialectes italiques : l'ombrien*. Paris, 1874. — M. Bréal. *Les tables eugubines*. Paris, 1875.

ractère : *arveitu*, *rere*, *runum*, correspondent au latin « advehito, dedit, donum ».

Nous nous contenterons de ces courtes indications sur les deux idiomes italiques frères du latin. En réalité ils ne diffèrent essentiellement de ce dernier, pas plus sans doute que ne diffèrent les dialectes grecs les uns des autres, mais beaucoup moins que les langues novo-latines ou les dialectes celtiques ne diffèrent entre eux.

Terminons par quelques mots sur les vieux alphabets italiques.

D'après Corssen (*op. cit.*, t. I. p. 1), ils descendraient de deux alphabets grecs. L'un de ceux-ci, le vieil alphabet dorien, ou quelque alphabet identique, aurait donné naissance à l'alphabet sabellique, à trois systèmes étrusques, à l'alphabet ombrien des tables eugubines, à l'alphabet osque que l'on trouve sur le cippe d'Abella. Ces différents systèmes possèdent tous, sauf le dernier, un double signe pour exprimer « s » : c'est le sigma grec, le sigma capital, figuré tel quel ou bien renversé à droite par un quart d'évolution de façon à figurer une sorte de M. Un alphabet dorien plus récent aurait donné naissance à l'alphabet falisque et à l'alphabet latin ; les plus anciens documents de ce dernier remontent à la fin du troisième siècle de notre ère. L'ancien *k* n'y était plus conservé que dans certains mots ; le *c* avait longtemps figuré le son *g* aussi bien que le son *k* et avait fini par être remplacé, pour le premier de ces offices, par un nouveau caractère, le G, procédant lui-même du C grâce à une minime modification. Du milieu du second siècle avant notre ère jusque vers le milieu du premier, c'est-à-dire pendant un espace d'environ cent ans, la règle parut s'introduire d'exprimer une voyelle longue en redoublant son caractère : on écrivit *aara*, *ree*, *Muucius*. Vers la même époque, un siècle environ avant notre ère, on figurait la voyelle longue *i* en lui donnant une forme plus haute

que celle des autres caractères du même mot : « dIvo, vIcus » ; parfois on se servit du signe I pour figurer la demi-voyelle « j » (notre « y »), comme dans « IUS, maIor ».

L'empereur Claude, au milieu du premier siècle de notre ère, prétendit doter l'alphabet latin de trois nouveaux signes. Pour distinguer la consonne *v* de la voyelle *u*, il proposa de représenter la première par le digamma grec renversé ; pour figurer les groupes *ps*, *bs* il proposa un C retourné, et enfin le signe ⊦ pour le son *ü*, notre « u » français, qui s'était introduit dans certains mots ; mais ces innovations n'eurent point un heureux succès et l'alphabet latin demeura ce qu'il était auparavant.

II. *Langues novo-latines.*

C'était au commencement de ce siècle une croyance fort répandue (et bien des personnes la conservent aujourd'hui encore) que le français provenait d'une langue *romane*, qui, vers la fin de l'empire et dans les premiers siècles du moyen âge, aurait succédé au latin, son ancêtre direct. Les écrits de l'illustre philologue Raynouard ne contribuèrent pas peu à propager cette théorie. On l'accepta volontiers ; on écrivit sur la langue romane, on commenta ses textes, et pour beaucoup de personnes le provençal actuel est encore cette langue romane. Raynouard s'était trompé et sa théorie devait disparaître peu après lui.

C'est qu'en effet il n'a point existé de langue romane : ce n'est pas à une langue romane que le latin a donné naissance, c'est à plusieurs langues romanes, à plusieurs langues novo-latines.

Il faudrait bien se garder, d'ailleurs, de ne voir dans ces nouveaux idiomes que du latin corrompu ; il n'en est

pas ainsi. Les langues novo-latines représentent tout autant de formes subséquentes du latin populaire parlé en Portugal, en Espagne, en France, chez les Grisons, en Italie et sur le bas Danube. A côté du latin littéraire, en effet, il existait une langue latine couramment parlée que les légionnaires et les colons apportèrent en Ibérie, dans les Gaules, en Dacie. C'est cette langue populaire qui se transforma et devint ici l'espagnol, ici le français, ici le roumain, de même qu'en Italie elle était devenue l'italien. Le latin littéraire, cependant, était de moins en moins intelligible pour le vulgaire et passait à la condition de langue ancienne, de langue classique, de langue morte.

« Quand le latin, écrit Littré, eut définitivement effacé les idiomes indigènes de l'Italie, de l'Espagne et de la Gaule, la langue littéraire devint une pour ces trois grands pays, mais le parler vulgaire, (j'entends le parler latin, puisqu'il n'en reste guère d'autre) y fut respectivement différent. Du moins c'est ce que témoignent les langues romanes par leur seule existence ; si le latin n'avait pas été parlé dans chaque pays d'une façon particulière, les idiomes sortis de ce parler latin que j'appellerai ici régional, n'auraient pas de caractères distinctifs, et ils se confondraient. Mais ces Italiens, ces Espagnols et ces Gaulois, conduits par le concours des circonstances à parler tous le latin, le parlaient chacun avec un mode d'articulation et d'euphonie qui leur était propre... Ces grandes localités qu'on nomme Italie, Espagne, Provence et France, mirent leur empreinte sur la langue, comme la mirent les localités plus petites qu'on nomme provinces. Et la diversité eut sa règle qui ne lui permit pas les écarts. Cette règle est dans la situation géographique, qui implique des différences essentielles et caractéristiques entre les populations. Le français, le plus éloigné du centre latin, fut celui qui l'altéra le

plus ; je parle uniquement de la forme, car le fond latin est aussi pur dans le français que dans les autres idiomes. Le provençal, que la haute barrière des Alpes place dans le régime gaulois du ciel et de la terre, mais qui les longe, est intermédiaire, plus près de la forme latine que le français, un peu moins près que l'espagnol. Celui-ci, qui borde la Méditerranée et que son ciel et sa terre rapprochent tant de l'Italie, s'en rapproche aussi par la langue. Enfin, l'italien, placé au centre même de la latinité, la reproduit avec le moins d'altération. Il y a dans cette théorie de la formation romane une contre-épreuve qui, comme toutes les autres épreuves, est décisive. En effet, si telle n'était la loi qui préside à la répartition géographique des langues romanes, on remarquerait çà et là des interruptions du type propre à chaque région, par exemple, des apparitions du type propre à une autre. Ainsi, dans le domaine français, au fond de la Neustrie ou de la Picardie, on rencontrerait des formations ou provençales, ou italiennes, ou espagnoles ; au fond de l'Espagne, on rencontrerait des formations françaises, provençales ou italiennes ; au fond de l'Italie, on rencontrerait des formations espagnoles, provençales ou françaises. Il n'en est rien ; le type régional, une fois commencé, ne subit plus aucune déviation, aucun retour vers les types d'une autre région ; tout s'y suit régulièrement selon les influences locales qu'on nommera diminutives en les comparant aux influences de région (1). »

Comme l'a fort exactement dit Ars. Darmesteter, le latin classique était une langue en quelque sorte artistique : le latin populaire c'est le latin vivant, le latin parlé, qu'est loin de représenter la latinité chrétienne. Quant au bas latin, c'est une langue nouvelle, c'est la langue populaire dans laquelle sont introduits des élé-

(1) *Dictionnaire de la langue française*, t. I, p. XLVII. Paris, 1863.

ments de la langue savante ; c'est la langue du moyen âge, la langue de saint Thomas : elle créera la langue française savante.

Cette origine latine des langues romanes est un fait acquis, démontré, éclatant que l'on ne saurait mettre aujourd'hui en question. La grammaire de Frédéric Diez, dont la première édition remonte à une cinquantaine d'années, a ruiné à jamais les théories ibériennes, celtiques ou autres qui se produisent encore de temps en temps.

Ce n'est pas à dire qu'il n'existe dans les langues novo-latines un fonds assez important de mots étrangers. Le français, par exemple, possède un certain nombre de mots d'origine celtique, tels que *arpent*, *lieue*, *dune*, *alouette ;* mais cette collection est loin d'être aussi considérable qu'on le peut supposer, et il est bon d'ajouter que tous ces mots, pour devenir français, ont dû auparavant se latiniser, et en somme c'est au latin que le français les emprunta. L'invasion des barbares apporta quelque chose comme quatre cents mots d'origine germanique, les relations avec l'Orient fournirent aussi leur contingent, mais la grammaire demeura essentiellement latine.

On cite comme langues novo-latines : le portugais, l'espagnol, le français, le provençal, l'italien, le ladin, le roumain.

Avant de parler de l'extension géographique de chacun de ces idiomes, et de dire quelques mots de leur physionomie particulière, nous devons tourner notre attention sur deux faits capitaux qui dominent toute cette étude. L'un de ces faits est le rôle de l'accent tonique dans la formation des mots novo-latins, l'autre est le passage de la déclinaison latine à l'état analytique des langues romanes.

On peut dire d'une façon générale, pour toutes les lan-

gues romanes, que la formation même de leurs mots est fondée sur la persistance de l'accent tonique (1) : là où était l'accent latin, là se trouvent l'accent italien, l'accent français.

Tel est le principe.

Des lois accessoires se sont jointes à ce principe, mais ne l'ont pas fait fléchir. Prenons pour exemple ce qui se passe dans la langue française.

A côté de la persistance de l'accent latin, le français nous montre deux principes accessoires : l'un est la suppression des voyelles brèves non accentuées qui précèdent la syllabe sur laquelle se trouve l'accent tonique ; l'autre est la chute de certaines consonnes médianes (2). L'accent, par exemple, est sur la voyelle « a » dans les mots *bonitatem*, *liberare*, *sanitatem ;* il reste sur la voyelle correspondante dans *bonté*, *livrer*, *santé*, et nous voyons que dans ces trois exemples, la voyelle inaccentuée « i » ou « e » a disparu. Dans *lier*, *douer*, la consonne médiane de *ligare*, *dotare* a également disparu.

Remarquons-le aussi, le français sacrifie aussi tout ce qui suit la syllabe accentuée ; les terminaisons masculines « essaim, peuplé, hôtel » portent toutes l'accent, et dans ses terminaisons dites féminines, « meuble, esclandre », il faut encore reconnaître que l'accent tonique est sur la dernière syllabe (dans le cas présent sur « eu, an »), car la voyelle terminale « e » n'est point prononcée et n'existe dans la poésie que d'une façon artificielle. En réalité, « esclandre, semaine » sont des mots de deux syllabes portant l'accent tonique sur la dernière de ces syllabes, sur « an » et sur « ai ».

(1) Littré. *Histoire de la langue française*. 6e édit., t. I, p. 242. Paris, 1873. — G. Paris. *Etude sur le rôle de l'accent latin dans la langue française*. Paris, 1862.

(2) Brachet. *Grammaire historique de la langue française*, Introduction, sect. II. Paris. — Scheler, *Exposé des lois qui régissent la transformation française des mots latins*. Bruxelles, 1875.

Mais il arriva un jour dans l'histoire de la langue française où le lexique tiré directement de la langue latine parut ne plus suffire, et l'on jugea bon d'accroître ce même lexique en empruntant au latin tels de ses mots qui n'avaient pas toujours leurs correspondants en français. On se contenta alors de calquer sur les mots latins les nouveaux mots demandés ; mais l'on ne pensa pas à observer cette loi fondamentale de la persistance de l'accent tonique, non plus qu'à faire tomber telle ou telle consonne médiane, telle ou telle voyelle inaccentuée. A ces nouveaux termes on a donné, ce qui pourrait sembler une sorte de dérision, le nom de « mots savants » ; aux mots d'origine vraiment naturelle, aux vrais mots français, aux mots corrects et bien formés, on a donné le nom de « mots populaires ». On ne s'en est pas tenu dans la fabrication des mots savants à calquer sur le latin des expressions dont le besoin se faisait sentir, on a reproduit également une foule de mots qui avaient donné déjà une forme populaire, une forme correcte, une vraie forme française. L'accent, par exemple, est sur la première syllabe dans les mots latins *debitum*, *cancer*, et, en français, ces deux mots étaient changés très régulièrement en *dette*, *chancre* : la formation dite savante les reprit, et, négligeant l'accent tonique, elle fabriqua les formes vraiment barbares de *débit*, *cancer*. Les mots *opérer*, *cumuler*, *séparer*, et une foule d'autres mots ont bien l'accent tonique sur la même syllabe que leurs modèles latins *operare*, *cumulare*, *separare* ; mais ce ne sont encore que des formes savantes, des formes secondaires, en face de *sevrer*, *combler*, *ouvrer*, qui ont négligé, comme il le fallait, la voyelle atone de la syllabe située avant la voyelle accentuée. De même encore les mots *lier*, *douer* représentent exactement le latin *ligare*, *dotare*, dont les mots savants *liguer*, *doter*, qui ont conservé la consonne médiane, ne sont qu'une

imitation arbitraire. On donne le nom de *doublets* aux termes d'origine savante et à ceux d'origine populaire qui proviennent d'un seul et même mot. Parfois même la forme populaire a disparu et la forme savante a seule persisté ; tel est le cas pour *facile*, *débile*, qui ne respectent point l'accentuation latine.

Arrivons au second fait capital, et non moins intéressant, qui domine, lui aussi, l'étude des langues romanes. C'est, avons-nous dit, le passage de l'état synthétique du latin qui possède une déclinaison de plusieurs cas, à l'état analytique des langues novo-latines qui ont perdu toute trace de déclinaison.

Dans les plus anciens monuments de l'italien et de l'espagnol, nous ne trouvons qu'une langue analytique, complètement analytique. Il n'en est pas de même de l'ancienne langue française ni de l'ancienne langue provençale ; à une certaine époque le français et le provençal se présentent, non pas avec des traces de cas, mais avec des cas véritables, avec deux cas, un cas sujet et un cas régime.

« Au moment, dit E. Littré, où une langue moderne se préparait dans les Gaules, le latin qu'on y parlait se présentait, quant à sa riche déclinaison, dans un état singulier : il employait assez bien le nominatif ; mais il confondait les autres cas et usait indistinctement de l'un pour l'autre ; c'est du moins ce qu'on trouve dans les monuments de l'époque, tout hérissés de ces solécismes. La langue nouvelle qui était en germe, ayant son instinct, porta la régularité dans ce chaos ; elle garda le nominatif, et des autres cas fit un seul cas, qui fut le régime. Aussi le français, dans sa constitution primitive, n'est point une langue analytique comme le français moderne ou comme le sont l'espagnol et l'italien dans leurs plus vieux textes ; il a un caractère synthétique, par conséquent plus ancien, exprimant les rap-

ports des noms entre eux et avec les verbes, non par des prépositions, mais par des cas (je me sers de ces termes synthétiques et analytiques pour dire que le latin exprime par des désinences significatives plus de rapports que ne le fait le français, qui, lui aussi, à bien des égards, demeure synthétique). C'est, comme on le voit, une syntaxe de demi-latinité, syntaxe qu'il a en commun avec le provençal. De sorte que les deux langues des Gaules, c'est-à-dire le français et le provençal, étant l'une et l'autre des langues à deux cas, se ressemblent plus entre elles qu'elles ne ressemblent à l'italien et à l'espagnol, qui, n'ayant point de cas, se ressemblent plus qu'ils ne ressemblent à la langue d'oïl et à la langue d'oc.

« Etre ainsi une langue à deux cas et retenir comme héritage du latin une syntaxe demi-synthétique ne fut pas dans le français une condition fugitive, qui n'ait laissé de trace que pour la curiosité de l'érudition. L'emploi en dura trois siècles. On ne parla et on n'écrivit que d'après cette syntaxe dans les onzième, douzième et treizième siècles. Le latin, qui est pour nous langue classique, reçoit beaucoup de louanges à cause de la manière dont sa déclinaison fait procéder la pensée. Je n'examine point la supériorité des langues à cas ou des langues sans cas : mais une part de ces louanges doit rejaillir sur l'ancien français, dont la déclinaison est amoindrie, mais réelle, et qui, à ce titre, est du latin au petit pied. Si le latin est, comme on le nomme souvent, une langue savante, l'ancien français réclame une part dans cette qualification ; et ceux qui ont traité de jargon notre vieille langue parlaient sans avoir aucune idée de ce qu'elle était. » (*Op. cit.*, *ibid.*)

La déclinaison du vieux français est fort simple. S'agit-il d'une forme répondant à la déclinaison latine en *us*, comme « dominus », le cas sujet du singulier prend

une *s*, qui n'est autre que l'ancien signe latin de ce même nominatif ; le cas régime du pluriel se suffixe également *s*, en souvenir du cas latin correspondant, « dominos ». Quant aux deux autres formes, le nominatif du pluriel et l'accusatif du singulier, elles restent telles quelles. C'est ce que montre, d'ailleurs, le tableau suivant :

Singulier.	Nominatif : *li chevals.*
	Accusatif : *le cheval.*
Pluriel.	Nominatif : *li cheval.*
	Accusatif : *les chevals.*

Nous sortirions de notre cadre en nous étendant sur ce sujet ; nous n'avons pas à faire ici l'histoire de la déclinaison de la langue d'oïl et de celle de la langue d'oc. Il nous suffit de constater qu'il y eut dans ces deux langues une période de véritable déclinaison que l'on ne peut retrouver dans les plus anciens textes des autres langues romanes. Ainsi que le dit E. Littré, il ne peut donc être question d'une vieille langue espagnole, d'une vieille langue italienne, au même sens qu'il peut être question d'une vieille langue française, d'une vieille langue provençale.

Cela dit, nous pouvons jeter un rapide coup d'œil sur chacun des sept rameaux qui forment la famille linguistique novo-latine.

1. La langue française. — Dès le premier siècle de notre ère les idiomes celtiques étaient supplantés en Gaule par le latin vulgaire ; il y eut à cela des causes nombreuses, des causes irrésistibles : au premier rang l'intérêt puissant qu'avaient les Gaulois à s'assimiler à leurs vainqueurs. La langue littéraire s'introduisit rapidement, elle aussi, et les écoles gauloises, formées sous la culture latine, eurent leur célébrité bien acquise. Cependant le latin vulgaire contribuait seul au développement de la langue populaire : c'était seulement au

latin vulgaire qu'elle allait devoir son origine. La langue classique disait, par exemple, « urbs, iter, osculari, os, hebdomas », mais le français prenait ses mots de *ville*, *voyage*, *baiser*, *bouche*, *semaine* aux formes populaires « villa, viaticum, basiare, bucca, septima ». Le nom de la langue française, de la langue d'oïl, était alors celui de « langue romaine rustique », et au huitième siècle les gens d'église prêchaient le peuple en langue rustique, en français. Une glose récemment découverte à Reichenau, et qui remonte à cette époque, est le plus ancien texte français que l'on connaisse. Les onzième, douzième et treizième siècles sont l'âge d'or de la langue d'oïl. « Alors se développent, dit A. Brachet, une littérature poétique pleinement originale, une poésie lyrique gracieuse ou brillante, une poésie épique grandiose, et dont la chanson de Roland reste l'expression la plus parfaite. L'Allemagne, l'Italie, l'Espagne s'approprient nos poètes et nos romans, les traduisent ou les imitent... Le douzième siècle, au moyen âge, le dix-huitième siècle dans les temps modernes, seront les principaux et les meilleurs représentants de notre génie national. » (*Op. cit.*, *ibid.*)

Nous avons parlé ci-dessus de la déclinaison à deux cas de la langue d'oïl : au quatorzième siècle elle disparaît, et le français du siècle suivant est décidément une langue moderne, une langue tout analytique, comme l'italien, comme l'espagnol.

Dès les premiers temps où nous puissions l'observer, la conjugaison française nous apparaît entièrement analytique. A côté des temps empruntés au latin, comme le présent *j'aime*, il s'en forme de nouveaux par le procédé moderne : *j'ai aimé*, *j'avais aimé*. Telle est l'origine du futur : *aimerai* est pour *aimer ai ;* les vieilles formes provençales et espagnoles ne laissent subsister aucun doute sur ce fait .Le latin, d'ailleurs, le latin classique

même, connaissait cette formation analytique d'un futur, et l'on trouve dans de bons auteurs « dicere habeo ». Quant au conditionnel, *j'aimerais*, ce n'est qu'une forme factice, calquée, en quelque sorte, sur le nouveau futur.

La langue française du moyen âge comptait un certain nombre de dialectes, indépendants les uns des autres et possédant leur littérature propre. Il n'en pouvait être autrement sous le régime de la féodalité. Les différences dialectiques n'étaient, pour la plus grande partie, que des différences d'ordre phonétique. Le bourguignon, le picard, le normand, durent céder toutefois devant le dialecte de l'Ile-de-France lorsque la famille des Capets fixa définitivement à Paris le centre du pays. Les dialectes descendirent alors, peu à peu, à la condition du patois : « C'est ainsi que le dialecte picard, le normand et le bourguignon furent en moins de trois siècles supplantés par le dialecte de l'Ile-de-France et tombèrent à l'état de patois, dans lesquels une étude attentive reconnaît encore aujourd'hui les caractères qui nous offrent les anciens dialectes avant les œuvres littéraires du moyen âge. Les patois ne sont donc pas, comme on le croit communément, du français littéraire corrompu dans la bouche des paysans ; ce sont les débris des anciens dialectes provinciaux, que les événements politiques ont fait déchoir du rang de langues officielles littéraires à celui de langues purement parlées. » (Brachet, *op. cit.*, p. 47.)

Le dialecte wallon conserva plus longtemps son individualité. Il comprenait deux variétés, le wallon liégeois et le wallon namurois (1) ; on l'a rattaché, mais à tort, au dialecte picard, dont il est bien distinct. Aujourd'hui, d'ailleurs, il n'est plus qu'un patois comme les autres dialectes du moyen âge, et le français littéraire l'a définitivement supplanté en tant que langue cultivée.

(1) Chavée. *Français et wallon*. Paris, 1857.

Nous avons eu l'occasion de parler incidemment, à plusieurs reprises déjà, des limites actuelles de la langue française. Au nord elle côtoie le flamand, un peu au-dessus de Calais, Saint-Omer, Armentières, Tourcoing, Ath, Liége, Verviers ; à l'est elle confine à l'allemand, en comprenant Verviers, Longwy, Metz, Dieuze, Saint-Dié, Belfort, Delémont, Fribourg, Sion ; plus au sud, à l'italien. Au midi de la France, enfin, elle s'étend sur le territoire tout entier des dialectes provençaux, dont nous allons nous occuper.

En Suisse, près de six cent mille individus ont le français pour langue maternelle (cantons de Neufchâtel, de Genève, de Vaud, majeure partie de Fribourg et du Valais, un cinquième de Berne) ; en Belgique, plus de deux millions, toute la partie sud-orientale de ce pays ; en Allemagne, plus de deux cent mille (Malmédy, Metz, Château-Salins). On parle également français dans les colonies anglaises de Maurice et du Canada.

2. Le provençal. — Faut-il admettre avec quelques auteurs que le français, ou langue d'oïl, et le provençal, ou langue d'oc, ne proviennent du latin vulgaire qu'indirectement et par l'intermédiaire d'une forme commune qui leur aurait donné naissance à l'un et à l'autre ? Cette opinion aurait besoin d'être justifiée ; jusqu'à ce jour elle ne repose que sur des assertions pures et simples. Disons qu'elle nous semble peu vraisemblable. Le latin populaire n'a pas dû se modifier uniformément par toute la Gaule ; qu'il n'ait pris, même, sur ce vaste territoire que deux sortes de physionomies distinctes, qu'il ne se soit transformé qu'en langue d'oïl et en langue d'oc, c'est ce qui peut à bon droit nous surprendre. Jusqu'à preuve contraire, il est sage, nous semble-t-il, de douter qu'un idiome commun franco-provençal ait jamais été parlé. La langue d'oïl et la langue d'oc se ressemblent sans doute d'autant plus que leurs textes

sont plus anciens, mais cela doit tenir uniquement au fait que plus elles sont anciennes, plus elles se rapprochent de leur origine commune.

Le provençal, ainsi que nous l'avons dit, eut comme la langue d'oïl une période semi-analytique durant laquelle il posséda une déclinaison à deux cas, un cas sujet, un cas régime. Nous avons dit tout à l'heure ce qu'il fallait entendre par là et nous pensons qu'il n'est pas utile d'y revenir. Quant à la conjugaison, elle est tout analytique comme dans la langue d'oïl ; c'est chez elle que l'on trouve cette ancienne forme du futur *dir vos ai*, « je vous dirai », qui montre bien clairement le mécanisme de la conjugaison nouvelle.

C'est un mot dont le sens est bien étendu que celui de provençal ; on prend ici une partie pour le tout. L'idiome de la Provence n'était et n'est encore, en effet, qu'un des dialectes de la langue d'oc : il faut ranger à côté de lui le languedocien proprement dit, le limousin, l'auvergnat, le dialecte du Dauphiné, le gascon.

On s'est demandé souvent si le *catalan*, qui occupe aujourd'hui en France une partie du département des Pyrénées-Orientales (la ligne frontière passe, en France, au nord de Perillos, Tautavel, Estagel, Montner, Neffiach, Ille, Rodès, Vinça, Marcevols, Arboussols, Molitg, Mosset, Puig Valador, Ruytor, Porté ; en Espagne, la Catalogne, Valence, les îles Baléares et qui s'étendait jadis sur le territoire aragonais), est un dialecte de la langue d'oc ou doit être considéré comme une branche distincte dans le système des langues novo-latines. Diez se prononce pour la première opinion ; mais la seconde peut être parfaitement défendue, et il ne serait pas étonnant qu'elle prévalût.

Les douzième et treizième siècle furent l'âge d'or de la littérature provençale, mais ses plus anciens monuments remontent à une époque antérieure ; la défaite des Albi-

geois lui porta un coup funeste : le français envahit peu à peu jusqu'aux Pyrénées toute la région où l'on ne parlait auparavant que la langue d'oc, et les dialectes de la France du Sud en sont arrivés à la condition du patois.

La limite actuelle des patois provençaux et français n'est pas très exactement fixée. On donne comme frontière extrême de la langue d'oïl Royan, Chalais, Ruffec, Montmorillon, Sainte-Sévère, Moulins, Charolles, Dôle, Vesoul, Belfort ; au sud de cette région commencerait la langue d'oc, dont les localités importantes situées le plus au nord seraient Bordeaux, Libourne, Confolens, Bellac, Boussac, La Palisse, Villefranche, Mâcon, Lons-le-Saunier, Poligny, Besançon, Montbéliard, Neuchâtel.

3. La langue italienne. — Telle que nous la connaissons, et dès ses plus anciens monuments, la langue italienne est incontestablement (et la raison en est naturelle) la mieux conservée des langues novo-latines, tant par son lexique que par ses formes elles-mêmes. Diez pense qu'il n'y a pas la dixième partie de son vocabulaire que l'on puisse faire remonter à une origine autre que la langue latine. Ce serait là un fait très remarquable. En tout cas, l'italien contient, sans nul doute, bien moins de mots allemands que n'en contient le français.

Au dixième siècle, on parlait déjà l'italien. Le latin vulgaire s'était assez transformé à cette époque pour que l'on pût, en Italie, lui donner le nom d'*italien ;* mais les monuments écrits de ce nouvel idiome ne remontent pas plus haut que le douzième siècle. C'est seulement au siècle suivant que naquit en Toscane la langue de la littérature italienne, langue purement littéraire, qui ne fut jamais parlée.

Quoi qu'il en soit, l'italien de cette époque avait la même physionomie que l'italien actuel, et il n'y eut

point une vieille langue italienne, comme il y a une vieille langue française, une vieille langue provençale.

L'italien compte un très grand nombre de dialectes, ce qui s'explique bien par la conformation même du pays où il est parlé. Ces dialectes sont fort nettement caractérisés. Dante, dans son traité « De vulgari eloquio », en énumérait quatorze : il les divisait en dialectes orientaux et dialectes occidentaux, ou, si l'on préfère, cisapennins et transapennins. Cette division a été remplacée avantageusement par celle de dialectes de la haute Italie, dialectes de l'Italie centrale, dialectes de la basse Italie. Dans cette dernière classe on range le napolitain, le calabrais, le sicilien, le sarde ; dans la seconde, le toscan, le romain, le corse ; dans la classe des dialectes du nord, le génois, le piémontais, les dialectes lombards et ceux de l'Emilie, et le vénitien. Chacun de ces dialectes possède une riche littérature ; un grand nombre d'entre eux ont des monuments qui remontent à l'époque de la Renaissance ; il y en a de plus anciens encore, par exemple le napolitain et le sarde.

La langue italienne actuelle franchit au nord les limites politiques de l'Italie.

En Suisse, cent trente mille individus parlent italien dans le canton du Tessin et la partie sud-ouest des Grisons.

En Autriche, une partie du Tyrol méridional parle italien, ainsi qu'une petite bande de la côte occidentale de l'Istrie.

4. La langue ladine. — On lui a donné les noms de langue des Grisons, rétho-roman, roumonche, roumanche. Il semble préférable de l'appeler langue ladine avec I. Ascoli, qui lui a consacré un très important travail (1).

(1) *Archivio glottologico italiano*, vol. I. *Saggi ladini*. Rome, Turin, Florence, 1873.

La langue ladine comprend, d'après cet auteur, trois groupes distincts : à l'est le frioulan, parlé par plus de quatre cent cinquante mille individus, en Italie sur les rives du Tagliamento et en Autriche jusqu'à Goritz. Au centre le ladin est parlé dans deux îlots du Tyrol autrichien, à quelque distance des deux rives de l'Adige, par plus de quatre-vingt-dix mille individus. A l'ouest, sous le nom de roumanche, il s'étend en sens transverse sur la plus grande partie du canton suisse des Grisons, où il est parlé par près de quarante mille personnes, ce qui fait en tout près de cinq cent quatre-vingt mille individus.

Ce nombre relativement peu considérable n'enlève point au ladin son caractère de véritable langue. C'est à tort que l'on a rattaché son groupe central et son rameau de l'est, le frioulan, à la langue italienne. Ils en sont parfaitement distincts par le matériel et les lois de leur phonétique et se relient intimement entre eux par ces mêmes éléments.

La littérature de la branche occidentale, celle de la langue des Grisons, est peu développée ; son plus ancien monument est une version du Testament chrétien remontant au seizième siècle. Les plus anciens documents du frioulan remontent au douzième siècle ; ce sont des inscriptions, assez courtes, il est vrai, mais qui suffisent pour caractériser la langue de cette époque.

5. L'espagnol. — C'est dans sa phonétique et son matériel lexique, où l'on rencontre entre autres éléments un assez grand nombre de mots arabes, que l'espagnol s'éloigne le plus du latin ; s'agit-il de la formation même des mots, il a conservé une remarquable fidélité. Ses textes les plus anciens remontent au milieu du douzième siècle ; peu abondants encore à cette époque, ils deviennent au siècle suivant de plus en plus riches. Il existe toutefois des traces plus antiques de la langue

espagnole ; ce sont, notamment, des mots cités par Isidore de Séville, qui vivait au septième siècle.

Les limites actuelles de l'espagnol sont tracées à l'ouest par le portugais, dont nous parlerons tout à l'heure ; au nord par le basque, dont nous avons indiqué ci-dessus la frontière, p. 149 ; à l'est il ne s'étend qu'en tant que langue littéraire sur la Catalogne et Valence, où la langue populaire est le catalan, dont nous avons parlé en traitant du provençal. L'espagnol, d'ailleurs, a conquis l'Aragon, où se parlait aussi autrefois le dialecte catalan. Il fait reculer également la limite méridionale de la langue basque : on ne parle qu'espagnol à Vitoria, Estella, Pampelune, Navascues ; Bilbao, Aoiz se trouvent déjà dans une zône mixte. Le basque cède ainsi bien plus rapidement au sud des Pyrénées qu'il ne cède au nord : c'est qu'en Espagne, ainsi que nous l'avons dit plus haut, il se trouve directement aux prises avec une langue officielle, tandis qu'en France, avant d'avoir affaire au français lui-même, il est en contact direct avec un dialecte de la langue d'oc, le gascon, dont la vitalité propre est déjà fortement menacée. S'il confinait immédiatement à la langue française, il cèderait devant elle aussi rapidement qu'il le fait au sud des Pyrénées devant l'espagnol.

6. Le portugais. — Il est fort rapproché de l'espagnol mais on ne peut le considérer comme un dialecte de cette dernière langue. Le portugais et le galicien, parlé au nord-ouest de l'Espagne, forment à eux deux un rameau roman bien indépendant. Leurs plus anciens monuments sont moins vieux que ceux de la langue espagnole et ne dateraient que des dernières années du douzième siècle. Le fonds de mots arabes que l'on rencontre en espagnol est à peu près le même que celui que l'on rencontre en portugais, mais cette dernière langue possède un certain nombre de mots d'origine française

étrangers à l'espagnol. Ils sont dus, pense-t-on, à la fin du onzième siècle, au temps de la domination de Henri de Bourgogne.

Le portugais, en dehors de son territoire européen, est parlé dans certaines contrées de l'Afrique et de l'Amérique, notamment au Brésil.

7. Le roumain. — Le roumain a pour origine le latin introduit en Dacie par les soldats de Trajan aux premières années du second siècle de notre ère : « Les soldats romains libérés du service, dit M. Picot, obtenaient, en même temps que l'honesta missio, le jus connubii et le jus commercii, c'est-à-dire le droit de faire commerce avec les barbares et d'épouser des femmes de leur race. Séparés à tout jamais du sol natal, cantonnés depuis vingt-cinq ans dans la même garnison, les légionnaires s'attachaient au pays où ils avaient vécu et combattu, et profitaient des facilités que la loi leur offrait pour s'y établir définitivement. C'est ainsi que se formèrent, sur les rives du Danube, les premiers groupes de population romaine, et à ces anciens militaires se joignirent bientôt des colons venus des diverses provinces de l'empire, et surtout des barbares attirés par l'appât du commerce. Les colonies militaires étaient fort nombreuses en Dacie à l'époque où les Romains furent obligés de se retirer. Il est à croire que la population purement romaine suivit les légions sur la rive droite du Danube, tandis que les individus issus du mélange des vétérans avec les barbares restèrent dans le pays où ils étaient nés, conservant la langue des vainqueurs, qu'ils avaient adoptée, et que ce sont eux qui ont donné naissance aux Romains. »

Nous aurons à parler plus loin de l'ancienne langue dace, dont la position dans la famille des langues indo-européennes est loin d'être fixée. Il est vraisemblable que le roumain conserve dans son vocabulaire des restes

de cette ancienne langue ; mais ces restes, quels sont-ils ? On ne peut le déterminer. Il faudrait, avant tout, connaître de l'ancien dace plus que nous n'en connaissons, plus que nous n'en connaîtrons jamais peut-être. On a pu, d'ailleurs, dresser la liste des emprunts faits par le roumain aux langues slaves dans les temps historiques ; elle n'est pas sans importance. On peut compter également un certain nombre de mots d'origine grecque ou d'autre source encore.

Longtemps on a supposé que le roumain était une langue slave. Cette erreur était due non seulement au fait que le roumain possède, ainsi que nous venons de le dire, une certaine quantité de mots empruntés aux langues slaves, mais aussi à ce qu'il a été écrit jusqu'en ces derniers temps en caractères cyrilliens, c'est-à-dire avec l'alphabet employé par le russe, le serbe, le bulgare. En certains cas, cet alphabet offrait des ressources assez précieuses, mais, par contre, il présentait parfois de grands désavantages. On l'a enfin abandonné et l'alphabet latin est définitivement accepté. Plusieurs systèmes de transcription se sont trouvés en présence lorsqu'il a dû être question d'adopter les signes diacritiques nécessaires pour compléter l'alphabet latin ; l'accord n'a pu s'établir complètement, mais un jour ou l'autre on arrivera, sans aucun doute, à ce résultat très désirable (1).

Les voyelles latines, en passant par la bouche des populations de la Dacie, ont subi, comme l'a montré Mussafia (2), deux variations capitales. D'une part les voyelles *e*, *o* portant l'accent tonique se sont changées, en certains cas, en *ea*, *oa*, c'est-à-dire en diphthongues ; d'autre part beaucoup de voyelles ont pris un son très

(1) Picot. *La Société littéraire de Bucarest et l'orthographe de la langue roumaine*. Paris.
(2) *Zur rumœnischen vocalisation*. Vienne, 1868.

sourd et presque nasal. Ce double phénomène est une des caractéristiques les plus importantes de la langue roumaine.

Le roumain possède un article ; mais, comme le bulgare, comme l'albanais, au lieu de le placer devant le substantif, il le lui suffixe : *omul* « l'homme ». Cette concordance de trois idiomes différents, mais parlés dans une même aire géographique, est des plus étranges. Doit-on y voir la trace d'une langue plus ancienne, telle par exemple que l'ancien dace, qui aurait laissé cet héritage aux idiomes divers par lesquels elle fut remplacée dans ce régime ? Si le fait est possible, il n'est point prouvé, et le champ des hypothèses demeure toujours ouvert.

La langue roumaine est très homogène, plus homogène qu'aucune autre langue novo-latine. L'acception donnée à tel ou tel mot peut bien varier de contrée à contrée, mais cela ne constitue pas des divisions dialectiques. On ne peut guère citer comme dialecte véritable que le roumain de la Macédoine, de la Thessalie, de l'Epire, le macédo-roumain.

En dehors de cet îlot détaché, le roumain est remarquablement compact. Il forme une sorte de cercle irrégulier de plus de cent lieues de hauteur (du Dniester au Danube) sur plus de cent lieues de largeur (d'Arad aux bouches du Danube). Outre la Valachie et la Moldavie, c'est-à-dire la Roumanie proprement dite, il comprend la partie nord-est de la principauté serbe, le banat de Temesvar, une grande partie de la Hongrie de l'Est, la plus grande part de la Transylvanie, la Bukovine du Sud, la Bessarabie, le territoire des bouches du Danube. Plus de huit millions et demi d'individus parlent aujourd'hui le roumain, près de neuf millions peut-être, dont la moitié environ dans la Roumanie proprement dite. Leur véritable nom n'est point celui de Valaques ; ce

nom leur a été donné par les Allemands. Ils le repoussent avec raison, se donnent à eux-mêmes le nom de « Roumains » et appellent leur langue la « langue roumaine », soucieux, avant tout, de perpétuer le souvenir de leur origine.

Les plus anciens textes connus de la langue roumaine ne remontent qu'à la fin du seizième siècle.

§ 5. Branche celtique.

Peu de mots ont prêté autant que ceux de « celte » et de « celtique » à toute une suite de malentendus anthropologiques, ethnographiques et archéologiques. Dans cette confusion la théorie fallacieuse des langues et des races a trouvé son compte plus que partout ailleurs. Enfin la question semble élucidée.

César avait raison, lorsqu'au début de son livre il divisait la Gaule en trois régions : l'Aquitaine au sud, la Celtique au centre, la Belgique au nord. Parlant de cette classification, — qu'appuient d'ailleurs un grand nombre d'autres textes, — l'anthropologie a établi que les Auvergnats et les Bas-Bretons actuels étaient les principaux représentants français de l'ancienne race celtique petite et brune, race qui n'avait et n'a rien de commun avec la race voisine du nord-est, grande, blonde, aux yeux bleus, à la carnation molle, et qui peut recevoir les noms de galate, galle, wallonne, belge, kimrique. Cette dernière race a été appelée souvent race celtique, mais à tort, et, ainsi que l'a démontré pertinemment P. Broca dans un excellent écrit, elle n'eut jamais droit à ce nom (1).

(1) *La race celtique ancienne et moderne. Arvernes et Armoricains, Auvergnats et Bas-Bretons*, Revue d'Antropologie, t. II, p. 577. Du même auteur : *Nouvelles recherches sur l'anthropologie de la France en général et de la Basse-Bretagne en particulier*, Mémoires de la Société d'anthropologie de Paris, t. III, p. 147.

La confusion qui a trop longtemps obscurci ce sujet était due pour une bonne part au nom même de « langues celtiques » donné d'une façon par trop générale aux idiomes que parlaient et les Celtes et les Galates du nord-est. De ce que ces derniers employaient une langue dite « celtique », on en a fait des « Celtes » ; c'était toujours la confusion des langues et des races. Il y aurait eu tout autant de raison à donner aux langues celtiques le nom de « galates », et si on ne l'a point fait, cela tient, sans nul doute, à ce que les Celtes (petite race brune brachycéphale) avaient pénétré sur le territoire qui plus tard reçut le nom de Gaule, bien avant que les Galates, leurs alliés par la langue, mais non par la race, y arrivassent à leur tour.

Il faut bien admettre, pour expliquer ce fait avéré de deux races très dissemblables parlant des idiomes très rapprochés d'une même langue, que les deux races en question aient vécu, à un moment donné, en grande connexité. C'est là un fait qui se reproduit partout de nos jours : il n'y a point, par exemple, de race française, mais bien plusieurs races parlant le français ; point de race italienne, mais bien plusieurs races parlant l'italien ; point de race allemande, mais bien plusieurs races parlant l'allemand. L'on ne saurait dire d'une façon précise quelle fut la région où Galates et Celtes parlèrent, presque en communauté, les langues qui reçurent plus tard le nom de « celtiques », mais les raisons anthropologiques portent toutes à croire que les Celtes étaient venus du sud-est de l'Europe : nous avons émis ailleurs l'opinion hypothétique que ce pourrait bien avoir été de la région du Dnieper et du bas Danube (1).

Nous n'insisterons pas sur ce côté de la question dite « celtique » : sans même rechercher s'il est permis d'at-

(1) Bulletins de la Société d'anthropologie de Paris, 1874.

tribuer exclusivement soit aux Galates, soit aux Celtes telle ou telle des deux branches des langues celtiques, nous aborderons immédiatement la question purement linguistique, qui seule nous intéresse ici d'une façon directe.

Les langues celtiques se divisent en deux groupes distincts et parfaitement caractérisés. A l'un de ces groupes on a donné les noms d'HIBERNIEN, de GADHÉLIQUE, GAÉLIQUE, à l'autre les noms de BRETON, de KIMRIQUE. Nous nous arrêterons, pour suivre la coutume habituelle et éviter toute espèce de malentendu, aux noms de gaélique et de breton. Au surplus, nous ne sommes pas en mesure de prétendre qu'il n'y eut pas dans l'antiquité d'autres branches de la famille linguistique celtique que les branches gaélique et bretonne. Le fait même est vraisemblable, si nous admettons, ainsi que de raison, que les idiomes de cette famille se sont étendus très anciennement sur de vastes contrées. Il ne paraît pas impossible que l'on découvre un jour dans l'Europe centrale, peut-être dans la région du Danube, des documents capables de confirmer cette supposition (1) ; mais aujourd'hui ces documents font encore défaut et nous n'avons à parler que des deux groupes ci-dessus nommés.

Le groupe GAÉLIQUE ou gadhélique, ou hibernien, comprend trois idiomes : l'*irlandais*, l'*erse*, le *mannois*, tous trois fort rapprochés les uns des autres.

L'importance de l'*irlandais*, dans l'étude des langues celtiques est considérable, non-seulement sous le rapport de la plus grande conservation de l'idiome, mais aussi sous celui de la richesse littéraire. A la vérité ce n'est qu'une richesse relative, mais la littérature des autres langues celtiques est si peu développée ! Les plus

(1) *La question celtique*. Bulletins de la Société d'anthropologie de Paris, 1874, p. 705.

anciens documents irlandais consistent spécialement en des gloses plus ou moins étendues insérées dans des manuscrits latins, soit à la marge, soit entre les lignes, et remontant au huitième siècle.

On rapporte au cinquième siècle au moins — époque à laquelle l'écriture latine pénétra chez les Hiberniens et les Bretons — les vieilles inscriptions irlandaises en caractères appelés « ogham ». L'origine de ces signes est loin d'être éclaircie, et nous devons nous borner à en faire cette simple mention. Au moyen âge la littérature irlandaise atteignit son apogée ; il reste de cette époque nombre de chroniques et de récits, sans parler de traductions d'œuvres étrangères.

Au temps de la Renaissance, l'irlandais entra définitivement dans sa période d'extinction : à l'heure actuelle, d'après Ravenstein (*Journ, of the Statist. Soc.*. t, XXIII. partie III), on compte près de 820.000 individus parlant irlandais en Irlande, soit 15 sur 100, et environ 50.000 en Angleterre. L'anglais occupe la partie orientale de l'île, plus de la moitié.

Sa situation géographique a mieux conservé le celtique écossais — l'erse — des empiètements de la langue anglaise. En Ecosse, 300.000 individus parlent l'erse, soit 9 pour 100 de la population ; 8.000 le parlent en Angleterre. Le gaélique d'Ecosse occupe toute la région septentrionale du pays, sauf un petit territoire à l'extrême nord-est, et la partie centre et ouest : soit, approximativement, les contrées de Caithness du Sud, Sutherland, Inverness, Argyle, Perth occidental ; il s'étend également sur les îles avoisinantes et sur celles qui se rapprochent de l'Irlande ; mais, au nord, dans les Orcades et Shetland il est inconnu.

Si la littérature du gaélique d'Ecosse est moins ancienne que la littérature irlandaise, elle possède toutefois le grand avantage d'avoir plus fidèlement gardé la

mémoire des traditions anciennes. Les poèmes apocryphes d'Ossian, qui soulevèrent, il y a cent ans, tant de controverses, avaient sans aucun doute un fond de vérité, et, à cette heure encore, les montagnards écossais sont loin d'avoir oublié tous les récits de leurs ancêtres.

Le dialecte de l'île de Man n'offre qu'un intérêt secondaire. D'après quelques auteurs, il serait parlé par un tiers des habitants du pays ; il ne le serait, d'après d'autres personnes, que par un quart ou un cinquième de la population (*Transactions of the Philological Society*, 1875, p. 172).

Le groupe BRETON, ou kimrique, comprend le *gallois*, le *cornique*, le *breton*, le *gaulois* ; deux de ces idiomes sont éteints, les deux autres vivent encore.

C'est au *gallois* qu'appartient la plus vivace des littératures celtiques actuelles. Dès le huitième siècle on trouve quelques gloses en gallois, aussi anciennes, par conséquent, que les gloses irlandaises dont nous avons parlé ; il est vrai qu'elles sont beaucoup moins importantes sous tous les rapports. D'ailleurs, c'est au moyen âge que se place la belle époque de la littérature galloise, notamment aux onzième, douzième, treizième siècles, qui ont vu paraître nombre de chroniques et de poésies. Aux approches de la Renaissance, le gallois sembla fort en danger ; il a repris cependant une certaine vitalité et c'est encore une langue écrite.

On compte, dans le pays de Galles, 934.500 individus parlant gallois ; on en compte, en Angleterre, 62.000.

Quant au cornouaillais, ou *cornique*, il s'est éteint au siècle dernier. Le plus ancien monument de sa littérature — un glossaire qui porte le titre de *Vocabula britannica* — date du treizième ou peut-être du douzième siècle.

On peut attribuer à l'époque de la Renaissance quelques autres écrits en langue cornique, notamment une

sorte de mystère chrétien sur la Passion ; nombre de mots anglais y ont déjà pénétré.

Le *breton*, ou armoricain, n'offre pas de très anciens documents, et ceux que l'on attribue à une époque plus reculée que le quatorzième siècle ne remontent sans doute pas à cet âge. Le plus connu est la vie de sainte Nonne et de son fils. L'on ne peut dire d'ailleurs que la littérature bretonne soit absolument éteinte aujourd'hui; on récolte, au moins, tout ce qui reste des vieilles traditions des poésies anciennes et la publication de quelques pièces plus ou moins apocryphes ne doit point laisser mettre en doute l'authenticité d'un très grand nombre d'autres morceaux.

Le breton est parlé dans le département du Finistère et dans la partie occidentale des Côtes-du-Nord et du Morbihan ; il se divise en quatre dialectes, parmi lesquels celui de Léon est le mieux étudié. Le véritable domaine breton comprendrait 1.200.000 individus ; la zone mixte en comprendrait 47.000. (Sur la limite du breton et du français, voir *Revue celtique*, t. V, p. 278 ; *Revue d'ethnographie*, t. V, p. 1.) D'après Paul Sébillot se trouvent en pays de langue celtique : Plouha, Pléguien, Tressignaux, Kérouzern, Kergourio, Lanrodec, Saint-Fiacre, Vieux-Bourg, Saint-Mayeux, Saint-Connec, Naizin, Saint-Jean Brévelay, Sulniac, Ambon ; se trouvent en pays de langue française : Tréveneuc, Tréguidel,Plouagat, Saint-Laurent, Saint-Gildas, Saint-Gilles, Gueltas, Radenac, Trédion, Elven, Muzillac.

Les deux douzaines d'inscriptions que l'on possède de l'ancien *gaulois* ont été découvertes, pour la plupart, dans la région de la Saône moyenne : il y en a pourtant qui proviennent du Rhône méridional, de la Normandie orientale et encore d'autres contrées. Ecrites en caractères latins, parfois en caractères grecs — par exemple celle de Nîmes — les inscriptions gauloises n'ont pas

encore été expliquées, bien qu'elles aient donné lieu à des travaux d'une véritable valeur, ceux, par exemple, de Pichet (1). Mais il reste également des noms de lieux, des noms propres cités par les auteurs classiques, et tout cela est plus que suffisant pour permettre de classer l'ancien gaulois dans la branche celtique bretonne ; nous reviendrons tout à l'heure sur ce sujet.

On connaît l'expédition historique des Galates en Asie Mineure, où ils s'établirent. Leur langue, qui, d'après des témoignages anciens, était la même que celle des habitants de Trèves, disparut dès les premiers siècles de notre ère, au plus tard vers le quatrième. D'après G. Perrot, elle serait tombée en désuétude dans le courant du premier siècle (2).

Les langues celtiques ne possèdent pas, sans doute, comme les langues germaniques, une caractéristique de premier ordre, telle que l'est, chez ces dernières, la substitution des consonnes. Mais, tout en se liant très étroitement, tout aussi bien aux langues germaniques, d'un côté, qu'aux langues italiques, d'un autre côté, elles n'en présentent pas moins un caractère particulier très frappant.

On ne saurait, à la vérité, définir ce caractère d'une façon un peu précise, mais il résulte d'un ensemble parfaitement tranché. On peut dire, en principe, que toutes les langues celtiques, lorsqu'il s'est agi de la formation des mots, ont montré une forte tendance à la contraction.

Si nous jetons un coup d'œil sur l'ensemble des voyelles du vieil irlandais, nous constatons sans peine qu'il est fort rapproché de l'ensemble du vocalisme latin, et

(1) Revue archéologique, 1867, p. 272. Ibid. Alfred Maury, 1866, p. s. Whitley Stokes, *Gallische Inschriften*, Beiträge zur vergleichenden Sprachforschung, t. II, p. 100.

(2) *De la disparition de la langue gauloise en Galatie*, Revue celtique, t. I, p. 179. Paris, 1870.

ce que nous disons de l'ancien irlandais est applicable, d'ailleurs, non seulement aux autres idiomes gaéliques, mais encore aux idiomes du groupe breton.

En ce qui touche les consonnes, il existe aussi une bien grande ressemblance entre les deux branches de la famille celtique ; l'un et l'autre, par exemple, aspirent en certains cas les consonnes K, T, P de l'indo-européen commun. Mais ce fait est moins général en breton qu'en gaélique ; ainsi l'armoricain et le gallois disent *dec*, « dix », tandis que l'ancien irlandais disait *deich* (prononcez « deikh »). L'irlandais moderne, devenant de moins en moins correct, change ces *ch* en *g* ; ici, par exemple, il dit *déag*.

Il est d'ailleurs un fait très général et très caractéristique qui distingue en principe, sur ce terrain de la phonétique, le système gallique du système breton : c'est ce fait que les « k » de l'indo-européen commun persistent dans le groupe gaélique — sauf, parfois, leur changement en aspirée, ainsi que nous venons de le voir — tandis que dans le groupe breton, pour l'ordinaire, ils se changent en *p*.

Ce fait est très important, et nous devons en donner au moins un ou deux exemples : gallois *peduur*, *pedwar*, « quatre », armoricain *peuar*, *pewar* : le « k » primitif s'est changé en *p*. Il persiste dans la branche gaélique : irlandais *cethir* ; comparez le latin « quatuor », le lithuanien « keturi » ; gallois *pimp*, *pump*, « cinq » armoricain *pemp*, le « k » persiste en gaélique : irlandais ancien *cóic*, irlandais moderne *cuig*. Comparez le latin « quinque ».

Ce changement de « k » en *p* se présente très clairement dans l'ancien gaulois, et c'est là une des raisons qui font rattacher cet idiome au groupe breton. Nous savons, par exemple, que le *quinquefolium* latin, la « quintefeuille », portait en gaulois le nom de *pempe-*

dula ; comparez le gallois *pump*, « cinq », l'armoricain *pemp*. Et ce fait n'est pas isolé.

La déclinaison de l'irlandais a beaucoup souffert, en ce sens que les désinences qui avaient originairement la mission d'indiquer les différents cas, ont été presque toujours mutilées d'une façon grave ; parfois même elles sont tout à fait tombées, et l'on ne peut plus savoir à première vue si tel ou tel nom est à tel ou tel cas, plutôt qu'à tel ou tel autre. D'anciennes formes pronominales, se changeant en vrais articles ou prépositions, sont venues remédier à ce fâcheux état : ainsi la forme *athir* « père » ne dit rien par elle-même du cas où elle se trouve, mais *intathir* signifie le nominatif « pater » et *sinnathir* l'accusatif « patrem ».

L'on peut dire que la déclinaison est encore plus maltraitée dans le groupe breton ; c'est à peine s'il y reste trace des désinences indicatrices des cas, et l'article lui-même a perdu sa diversité. Ainsi, en armoricain, *roen*, « roi », signifie tout à la fois « rex, regem, regis », etc., et l'article *an* le précède toujours tel quel : *an roen*, « rex » ; *an roen*, « regis ». C'est uniquement aux prépositions, dont on usera selon les différents cas, qu'il appartiendra de déterminer la fonction du substantif ; nous ne suivons pas, en français, un autre procédé lorsque nous disons « *à* la femme, *de* la femme, *pour* la femme », etc. Tel est le dernier degré de l'analytisme.

La conjugaison gaélique et celle du groupe breton suivent essentiellement le même système, et l'on peut dire que ce système n'est pas sans offrir de grandes difficultés ; c'est là que réside la véritable peine de l'étude des langues celtiques. Ici encore, comme l'on pouvait s'y attendre, le groupe breton se montre bien moins conservé que le groupe gaélique.

Ce serait une entreprise à peu près interminable que celle de relever tout ce qui a été dit d'énormités sur les

langues celtiques. Il n'est pas rare, aujourd'hui encore, d'entendre expliquer le phénicien et l'étrusque par les racines celtiques ; il est moins rare encore d'entendre expliquer le basque par des mots bretons ou irlandais. Mais ce qui est bien plus fréquent, ce sont des théories quasi périodiques qui, en dépit de tout ce qui a été dit, écrit et mille fois prouvé relativement à l'origine des langues romanes, font encore dériver ces dernières langues des langues celtiques. Cette persistance de la celtomanie tient uniquement à ce que ceux qui en font profession ignorent trois choses capitales : les langues celtiques, le latin, les langues novo-latines. Il n'est pas un celtomane qui ne soit étymologiste ; l'étymologie est la condition vitale de la celtomanie.

Ce n'est pas à dire que les langues celtiques n'aient point fourni aux idiomes novo-latins un certain nombre de mots de leur vocabulaire, mais ce nombre est peu considérable. Ce sont des noms géographiques qui en forment le meilleur contingent ; tels les noms du Danube, des Alpes, des Ardennes. Les mots *lieue*, *dune*, *alouette* ont également une origine celtique ; mais cette origine n'est qu'indirecte, et pour passer du celtique au français, ces mots, comme nous l'avons dit en parlant des langues romanes, ont dû avant tout se latiniser.

§ 6. **Branche germanique.**

Il ne semble point que les noms de « Germain », de « Germanie », de « germanique », soient d'origine allemande ; on a cherché souvent à les expliquer, mais cela a toujours été en vain. Le véritable nom appelé à remplacer celui de « germanique » serait sans doute le nom de *tudesque*. Celui-ci n'est autre que l'allemand moderne

deutsch, l'ancien haut-allemand *diutisc*, correspondant à une forme plus ancienne *thiudisks*, adjectif dont le sens primitif était celui de « populaire, national ». Toutefois, le nom de « germanique » est trop accepté, trop usuel, pour que l'on puisse songer à le remplacer par un autre. Les Allemands eux-mêmes, qui protestent contre cette appellation et aiment à employer leur nom national, ne donnent point aux langues indo-européennes le nom d'indo-tudesques, mais bien celui d'indo-germaniques. C'est là une inconséquence assez singulière.

Le système des langues germaniques se divise en quatre branches distinctes : la branche gothique, la branche scandinave, la branche bas-allemande, la branche haut-allemande.

Avant de porter notre attention sur chacune de ces branches en particulier, nous devons jeter un coup d'œil sur l'ensemble du système.

La grande caractéristique du groupe germanique consiste dans la façon dont il traita les explosives de la langue commune indo-européenne « k, t, p ; g, d, b ; gh, dh, bh ». Il les renforça toutes. Là où l'indo-européen commun avait une explosive aspirée, il mit une explosive non aspirée ; là où l'indo-européen avait une explosive faible, il mit une explosive forte. Quant aux explosives fortes de l'indo-européen commun, il les changea en sifflantes : « k » devint *h*, « p » devint *f* et « t » fut remplacé par la sifflante *th* dur des Anglais (le *th* de « three, thank », non pas celui de « they, the »).

Lors donc que le sanskrit, fidèle aux explosives primitives, dit *bhrâtâ* « le frère », le gothique dit *brôthar*, changeant l'aspirée en non aspirée, et la forte en sifflante ; tandis que le sanskrit dit *ajras* « le champ », le grec ἀγρός, le latin *ager*, le gothique dit *akrs*, changeant l'explosive faible en explosive forte.

Rien de plus simple que cette loi. Ajoutons qu'elle est constante. Pour qu'elle ne s'applique pas, il faut qu'il se rencontre un empêchement physiologique, il faut que l'explosible, qui devait être renforcée, soit, par exemple, précédée de « s ». En ce cas elle demeure telle quelle; c'est ainsi qu'au sanskrit *asti* « il est », au lithuanien *esti*, correspond le gothique *ist*.

Cette loi du renforcement des explosives primitives, cette grande caractéristique du système germanique pris en général fut, par la suite des âges, étendue, complétée, mais elle demeura la base même du système tout entier.

En dehors des nouvelles sifflantes *f*, *h*, *th* dur, *th* doux et *z*, les vieilles langues germaniques n'ont guère augmenté le tableau des consonnes de la langue commune indo-européenne. Par contre, elles ont perdu les trois explosives aspirées « gh, dh, bh », dont elles ont fait, ainsi que nous venons de le dire, trois explosives simples. En ce qui concerne les voyelles, les idiomes germaniques sont moins purs ; ils les ont singulièrement modifiées et ils possèdent une grande richesse de diptthongues. Leur ancienne déclinaison, moins bien conservée que celle de la plupart des autres idiomes indo-européens, est cependant assez fidèle en bien des points ; mais la conjugaison a éprouvé des pertes considérables, celle, par exemple, de la plus grande partie des temps organiques.

I. *Gothique*.

Nous serions tenté, si l'usage ne s'y opposait, de ne pas écrire ce mot de *gothique* avec l'orthographe ordinaire et reçue, mais de nous conformer à l'orthographe plus correcte de « gotique ».

Les Goths, en effet, écrivaient leur nom avec un *t*,

non pas avec un *th*, ce qui était bien différent ; nous avons vu, en effet, que le *th* des vieilles langues germaniques était une véritable sifflante, non pas une explosive plus ou moins aspirée.

Les Romains écrivaient correctement « goticus », et c'est aux historiens grecs que nous devons la vicieuse orthographe du mot « gothique ».

On a cru longtemps que le gothique était le père commun de tous les idiomes germaniques; il n'en est rien. Plus correct que chacun d'eux dans son ensemble, plus rapproché de l'indo-européen commun, il cède le pas parfois à tel ou tel de ses congénères. Il faut le placer, en réalité, à côté du vieil islandais, à côté des vieux idiomes bas-allemands, et souvent aussi sur le même pied que l'ancien haut-allemand, bien que ce dernier, à un point de vue spécial, le cède beaucoup à tous ses alliés. Un grand nombre des formes du bas et du haut-allemand s'expliquent sans doute par le gothique, mais aucune d'elles n'en provient directement ; gothique, scandinave, bas et haut-allemand, remontent tous, en un mot, à une forme commune qu'aucun d'eux ne représente d'une façon complète.

A quelle époque et où fut parlée la mère commune des langues germaniques ? C'est ce qu'il sera peut-être impossible de jamais savoir.

Quant à la langue gothique, nous la connaissons sous la forme qu'elle avait au quatrième siècle de notre ère, par la traduction des Testaments juif et chrétien due à Vulfila — l'Ulphilas des auteurs grecs — évêque des Goths établis en Mésie. Elle devait s'éteindre au neuvième siècle, cinq cents ans plus tard.

Son système vocalique est le moins compliqué de tous ceux des anciennes langues germaniques.

Nous avons parlé du renforcement que l'ensemble des langues germaniques fit subir aux consonnes explosives

de l'indo-européen commun. Le gothique, après avoir appliqué cette loi rigoureuse, fit ensuite subir aux sifflantes obtenues par ce renforcement une modification nouvelle. Parfois, chez lui, *h* (produit d'un « k » plus ancien) devint *g* ; parfois *th* (produit d'un « t » plus ancien) devint *d* ; parfois *f* (produit d'un « p » plus ancien) devint *b*. Ce phénomène est très remarquable; les nombreux exemples qu'il fournit ont été cités maintes fois comme constituant tout autant d'exceptions au principe général du renforcement, mais, en fait, il n'en est point ainsi. Il y a là un phénomène accessoire, un phénomène secondaire, mais un phénomène distinct. Nous verrons, en parlant des idiomes bas-allemands, chez lesquels cette loi secondaire peut être prise en flagrande application, comment le nom de loi de polarité peut se justifier. Pour l'instant, nous ne faisons que constater ses effets dans la langue gothique, où elle est peu fréquente, mais où elle existe cependant.

Les lois phonétiques du gothique sont peu nombreuses, mais assez importantes. Une des plus caractéristiques est celle-ci : dans les mots de plus d'une syllabe les voyelles *a* et *i* précédant une consonne placée à la fin du mot, tombent.

Une autre loi phonétique importante est celle qui, en principe, change les *i* en *ai* et les *u* en *au* devant les consonnes *r* ou *h* ; c'est là une loi particulière au gothique.

Dans la déclinaison nominale le gothique a perdu toutes les formes du duel et le cas ablatif ; presque tous ses datifs sont empruntés au vocatif.

De la conjugaison organique il n'a gardé que le présent et l'ancien parfait redoublé, ce dernier au moins pour une partie de ses verbes ; il ne présente plus trace des deux aoristes, de l'imparfait, du futur. Il rend ce dernier temps par des formes du présent et s'est fabriqué

pour la plupart des verbes dérivés une sorte de parfait.

Le gothique disparut sans laisser de descendants. C'est ainsi qu'avaient également disparu bien d'autres idiomes germaniques parlés vers la même époque, et dont il ne nous est parvenu aucun document, ceux, par exemple, des Vandales, des Hérules, des Burgondes.

II. *Langues scandinaves.*

L'ancienne langue *nordique* fut portée en Islande par des colons norwégiens ; grâce au lent développement que reçut la civilisation dans cette île lointaine dont les communications avec le continent étaient difficiles, le vieux nordique s'y maintint bien plus aisément que dans les autres pays scandinaves.

L'islandais moderne diffère peu en réalité, de cette ancienne langue, et sa supériorité sur tous ses congénères européens, non-seulement sur les langues germaniques modernes, mais encore sur les idiomes slaves, celtiques, novo-latins et autres, serait bien assurée si le lithuanien n'existait pas. Le côté faible de l'islandais moderne, c'est d'avoir subi cette loi de renforcement dont nous avons parlé ci-dessus, mais qui était commune à toute la famille germanique et dont il ne pouvait s'affranchir.

La phonétique de l'ancien nordique est beaucoup plus délicate que celle du gothique. On y compte une vingtaine de voyelles différentes, longues ou brèves, et plusieurs diphthongues. Le nombre des consonnes est également d'une vingtaine. Outre les explosives fortes et faibles, le vieux nordique possède les deux sifflantes *f*, *h*, la sifflante *th* dur (de l'anglais « thick ») et sa correspondance douce (anglais « they, the »). Le germa-

nique se distingue, d'ailleurs, des autres langues germaniques par une plus grande disposition à l'assimilation des consonnes. En somme, la déclinaison est aussi bien conservée qu'en gothique et la conjugaison a éprouvé les mêmes pertes. Il s'y est formé par des moyens factices, par des moyens analytiques, un futur, un conditionnel et un nouveau parfait.

C'est en Islande que furent composés les plus beaux monuments de la vieille littérature nordique, les deux Eddas, recueil d'anciens récits mythologiques. La première est en vers et remonte au onzième siècle ; la seconde, plus récente d'une centaine d'années, dit-on, est en prose : c'est une sorte de complément de la première Edda.

Les langues scandinaves modernes sont au nombre de quatre : l'*islandais*, le *norwégien*, le *suédois*, le *danois*. Pour certains auteurs l'islandais seul descendrait directement de l'ancien nordique, et les trois autres idiomes scandinaves proviendraient de dialectes différents de cette ancienne langue, bien qu'étroitement alliés avec elle. Pour d'autres auteurs, les quatre langues scandinaves actuellement parlées auraient toutes le vieux nordique pour ancêtre direct. Ce qui semble incontestable, en tout cas, c'est la parenté plus intime de l'islandais et du norwégien, du suédois et du danois. On peut ainsi les diviser en deux groupes assez distincts (1). L'islandais et le norwégien, par exemple, gardent les diphthongues de l'ancienne langue, tandis que le suédois et le danois les changent en voyelles longues ; danois et suédois conservent au commencement des mots certains groupes de consonnes que l'islandais et le norwégien ont perdus ou ne prononcent plus entièrement.

Le norwégien, dont la littérature est uniquement populaire, a perdu beaucoup de terrain. Le suédois, au

(1) Th. Möbius. *Dänische formenlehre*, p. 2. Kiel, 1871.

contraire, en a gagné et possède une véritable littérature. Non seulement il occupe une partie de la péninsule scandinave, mais il s'étend sur les deux bandes de territoire du littoral finlandais : l'une de ces bandes donne sur le golfe de Bothnie, avec Vasa pour point central ; sa longueur est d'environ une cinquantaine de lieues, sa largeur est fort minime. L'autre bande, un peu plus importante, occupe la partie ouest de la rive septentrionale du golfe de Finlande, avec Helsingfors pour point central. Du côté de la terre ces deux territoires sont enveloppés par les régions où se parle le suomi.

On peut dire d'une façon générale que le suédois a conservé mieux que le danois la physionomie de l'ancien scandinave. Les consonnes *k*, *t*, *p*, par exemple, s'affaiblissent à la fin des mots dans la langue danoise en *g*, *d*, *b*, tandis qu'elles persistent en suédois. De toutes les langues scandinaves actuellement parlées, le danois est, en somme, la plus moderne, si l'on envisage ses formes. Il n'est point parlé seulement en Danemark : en Norvège on l'écrit couramment et il est parlé par les classes lettrées, tandis que le norwégien n'est, ainsi que nous l'avons dit, qu'un idiome populaire ; il s'étend également sur la partie nord du Slesvig et comprend la ville de Flensborg. Le danois se divise d'ailleurs en plusieurs dialectes. Ses plus anciens documents remontent au treizième siècle, mais sa forme actuelle semble être née au seizième siècle du dialecte séelandais. On y rencontre un certain nombre de mots étrangers, empruntés au latin, au suédois, au français et surtout à l'allemand.

III. *Bas-allemand.*

Cette branche des langues germaniques est divisée en un assez grand nombre de rameaux. Elle aurait donné naissance tout d'abord à deux tiges distinctes, celle du

saxon et celle du frison ; la première de ces tiges aurait produit ensuite, plus ou moins directement, une demi-douzaine d'idiomes.

On s'accorde, pour l'ordinaire, à dresser le tableau que voici :

Bas-allemand.	Saxon.	Anglo-saxon.	Anglais.
		Vieux saxon.	Bas-allemand proprement dit.
			Néerlandais. } Hollandais et flamand.
	Frison.		

Nous ne pouvons savoir directement ce qu'était la forme commune du bas-allemand, au sens général du mot, non plus que ce que pouvait être la forme commune saxonne d'où devaient sortir l'anglo-saxon et le vieux saxon. Quant à ces deux derniers idiomes, ce sont des langues historiques fort bien connues.

Le *vieux saxon* était parlé du Rhin à l'Elbe, au sud du frison, qui occupait les pays allemands du Nord. Nous possédons du vieux saxon un monument important, le poème chrétien du Sauveur (Hêliand), conservé dans deux manuscrits qui datent du neuvième siècle.

L'*Anglo-saxon* remonte au septième siècle, au moins en Angleterre, et lui aussi a son épopée, le poème de Beovulf, que l'on rapporte à cette ancienne époque.

La langue de ces deux vieux idiomes bas-allemands n'était pas fort différente, mais elle présentait cependant une certaine quantité de divergences très caractérisées, notamment dans la phonétique. Le système des voyelles du vieux saxon est de beaucoup plus simple que celui de l'anglo-saxon ; ce dernier est véritablement assez compliqué et son échelle vocalique est très chargée.

On divise l'anglo-saxon en deux périodes : l'une, celle de l'anglo-saxon proprement dit, ou du vieil anglo-saxon, s'étend jusqu'au commencement du douzième siècle; la seconde, celle du demi-saxon, jusqu'au milieu du treizième.

La période de l'ancien *anglais* est également longue d'un siècle, elle va de l'an 1250 à l'an 1350 environ ; avec elle commence une décadence rapide des formes de la langue : il ne reste plus des anciens cas que le seul génitif, et encore ce cas est-il remplacé maintes fois par l'emploi de prépositions. Au milieu du quatorzième siècle s'ouvre la période du moyen anglais, qui doit durer deux cents ans, et qui continue, d'une façon plus marquée encore, l'œuvre de la période précédente. Lorsqu'apparaît enfin le nouvel anglais, l'anglais moderne, au milieu du seizième siècle, la langue est presque entièrement analytique. Les dialectes sont nombreux, mais on peut dire qu'ils en sont tous arrivés au même point de simplification grammaticale. Au demeurant, il reste encore assez de grammaire chez eux tous, aussi bien que dans la langue littéraire, pour témoigner du caractère essentiellement germanique de l'anglais. D'ailleurs, l'introduction considérable de mots français dans la langue anglaise (1) n'affecte en rien sa grammaire, comme on a pu le croire et le dire ; l'anglais n'est point une langue mixte, c'est une langue foncièrement germanique, dont les formes ont eu à souffrir plus que celles de tous les idiomes de la même famille.

Nous revenons à la seconde branche du saxon, au vieux saxon. Son système de voyelles, avons-nous dit, était beaucoup plus simple que celui de l'anglo-saxon, il connaissait beaucoup moins de voyelles que ce dernier, et ses rejetons contemporains ont, eux aussi, une échelle vocalique beaucoup moins compliquée que ne l'est celle de l'anglais. Il a donné naissance à deux rameaux frères, le rameau du bas-allemand proprement dit et le rameau néerlandais.

(1) Thommerel. *Recherches sur la fusion du franco-normand et de l'anglo-saxon*. Paris, 1841.

Le *bas-allemand proprement dit*, ou « plattdeutsch », est la langue populaire des basses régions de l'Allemagne du Nord. Le bas-allemand a gagné considérablement vers l'est sur des régions où se parlaient autrefois des idiomes slaves et même des idiomes lettiques (le vieux prussien et le lithuanien), mais il ne s'est pas élevé à la condition de langue littéraire, et le haut-allemand moderne, l'allemand, a rendu pour jamais infructueux les essais que l'on pourrait faire en ce sens.

La seconde branche du vieux saxon est le néerlandais, qui se divise en deux langues fort rapprochées l'une de l'autre, presque identiques, le *hollandais* et le *flamand*. Ce dernier est souvent regardé comme un dialecte du hollandais ; rien n'est plus inexact. Flamand et hollandais doivent être placés sur un pied d'égalité, et ils sont si rapprochés l'un de l'autre, que l'on a pu dire, avec assez de raison, qu'il n'y avait entre eux qu'une simple différence de prononciation. Le flamand est parlé aujourd'hui par deux millions quatre cent et quelques milliers d'individus ; quant au nombre des personnes parlant hollandais, on peut l'évaluer approximativement au chiffre de trois millions et demi, ce qui donnerait comme nombre total des individus de langue néerlandaise le chiffre de six millions, y compris les Flamands français du département du Nord. Certains auteurs vont jusqu'au chiffre de six millions et demi.

La ligne qui sépare le français du flamand est, la plupart du temps, assez horizontale. Elle laisse au nord, en pays de langue flamande, Gravelines, Hazebrouck, Courtrai, Halle, Bruxelles, Louvain, Tongres ; au sud, en pays de langue française, Calais, Saint-Omer, Armentières, Tourcoing, Ath, Nivelles, Liège, Verviers.

Nous n'avons parlé jusqu'à présent que d'une seule des branches du bas-allemand, la branche saxonne. L'autre branche est incomparablement moins impor-

tante. Elle ne comprend que le *frison*. C'est une langue assez ancienne parlée sur la côte de la mer du Nord, aussi bien sur le continent même que dans les îles qui lui font face. Les Frisons sont loin d'avoir entrepris les émigrations devant lesquelles les autres Bas-Allemands ne reculèrent point ; ils demeurèrent dans les régions où ils étaient établis, et leur langue conserva certains caractères de grande antiquité, malgré l'influence que purent exercer sur elle les idiomes avoisinants, le néerlandais, l'allemand, le danois.

Depuis longtemps le frison n'est plus une langue littéraire ; l'allemand l'a totalement relégué au second rang, comme il a fait, d'ailleurs, au bas-allemand proprement dit.

Plus haut, en parlant de la langue gothique, nous avons dit quelques mots d'un principe de la phonétique germanique qui fut secondaire à la loi générale du renforcement des explosives et qui trouva son application dans les quatre branches de cette famille. Nulle part, avons-nous ajouté, ce nouveau phénomène n'est plus aisément surpris en application flagrante que dans les idiomes qui appartiennent au groupe bas-allemand. C'est ce que nous allons montrer.

Nous savons que, d'après le principe du renforcement des explosives organiques, les « k, t, p » de la langue indo-européenne commune étaient devenus, dans le système germanique général, de véritables sifflantes : *h*, *th*, *f*. Le nouveau phénomène dont nous avons à parler consiste en ceci : les trois sifflantes dont il est question ne se maintiennent pas toujours telles quelles, parfois elles se changèrent en *g*, en *d*, en *b*, et cela dans tous les idiomes germaniques.

Ce changement n'eut pas lieu directement ; il y eut un intermédiaire entre *h* et *g*, un intermédiaire entre *th* et *d*, un intermédiaire entre *f* et *b*.

C'est ici que les idiomes bas-allemands nous sont d'une utilité capitale : ils nous présentent maintes fois, en effet, coexistant à côté l'un de l'autre, les différents termes de cette série. Nous savons, grâce à eux, que l'intermédiaire entre la sifflante dure et l'explosive faible fut la sifflante douce correspondante. Ainsi c'est par l'intermédiaire de *v* que l'on passa de *f* à *b*, c'est par l'intermédiaire d'un *h* doux que l'on passa de *h* dur à *g*, c'est par l'intermédiaire de *th* doux (celui de l'anglais « they, there ») que l'on passa du *th* dur (de l'anglais « thick, thirst ») à l'explosive faible *d*. Cela fit donc trois degrés successifs.

Un ou deux exemples rendront la chose plus compréhensible. Tandis que le gothique sous la forme *tha-* reproduit le pronom organique « ta », en changeant l'explosive forte en une sifflante *th* dur, l'anglais, dans son article *the*, change la sifflante dure en sifflante douce, et le hollandais, poussant l'évolution jusqu'au bout, dit *de*, tout comme le flamand. C'est ainsi encore que le hollandais *doorn* correspond au gothique *thaurnus* « épine », *voor* à *faur* « pour », *vol* à *fulls* « plein ». Ce n'est pas à dire, d'ailleurs, que l'anglais en reste toujours à la consonne intermédiaire, ce n'est pas à dire, non plus, qu'il atteigne toujours cette même consonne ; assurément non. Les *th* durs qu'il présente si souvent le montrent encore dans la première période : *thorn* « épine », par exemple, est avec le gothique au premier degré, tandis que le *doorn* hollandais est déjà au troisième. Mais cela ne fait rien à l'affaire. Nous pouvons prévoir en toute sûreté le temps où tous les *th* anglais se seront changés en *d*, comme ç'a été le cas pour le hollandais et le flamand. Un certain nombre des dialectes anglais en sont déjà arrivés à cette troisième période ; dans les pays de Kent et de Sussex, par exemple, l'on dit *dey* « eux », *de* « le » pour « they, the » de la langue litté-

raire ; à Wight l'on dit *vor* pour *for* « pour » ; dans les pays de Dorset, de Devon, de Somerset, même changement de *f* en *v* (1). La langue littéraire devra faire à son tour le chemin que parcourent aujourd'hui les dialectes, et elle arrivera à l'état qui caractérise à présent le flammand et le hollandais.

IV. *Haut-allemand.*

Le haut-allemand moderne, l'allemand, occupe au centre de l'Europe un territoire assez considérable. Au nord il est la langue littéraire, la langue cultivée des pays où se parle le bas-allemand proprement dit, le platt-deutsch, et il s'étend en cette qualité jusque près de Flensborg dans le Slesvig méridional. Au nord-est, il atteint presque la frontière russe, que longe encore cependant, derrière Memel et Tilsit, une petite bande de pays lithuanien. Une bande plus considérable de territoire polonais l'empêche de confiner à la frontière du royaume de Pologne, mais il occupe, du moins, toutes les localités importantes de cette région : Graudenz, Thorn, Posen, Oppeln. Contournant de l'est à l'ouest le nord du pays tchèque et redescendant par les environs de Pilsen et de Budweis vers Brünn, en Moravie, la frontière de l'allemand gagne Presbourg, longe sur une quarantaine de lieues le pays de langue magyare et englobe la Styrie du Nord (Gratz), la Carinthie septentrionale (Klagenfurt), la plus grande partie du Tyrol, les trois quarts de la Suisse (Coire, Altorf, Brieg, Laupen, Soleure, Bâle). Laissant Belfort vers l'ouest, elle remonte les Vosges jusqu'à la hauteur de Strasbourg et oblique vers le nord-ouest en englobant Thionville et Arlon ; gagnant Aix-la-Chapelle, elle suit dès lors la frontière

(1) C.-Fr. Koch, *Historische grammatik der englischen sprache*, t. I, p. 27. Weimar, 1863.

néerlandaise. Dans cette énumération l'Autriche-Hongrie entre pour 9 millions d'individus, la Suisse pour 1.755.000.

C'est du seizième siècle que date l'allemand moderne. La branche germanique dont il est aujourd'hui le représentant avait passé auparavant par deux périodes, d'abord par celle du vieux haut-allemand, puis par celle du moyen haut-allemand. Il nous reste, pour terminer avec les langues germaniques, à parler de ces trois périodes.

Il y a deux sortes de haut-allemand, le haut-allemand rigoureux et le haut-allemand qui ne s'est pas soumis à la loi commune. Ce ne sont point là deux langues ; en fait, il n'y a qu'une seule et même langue allemande, mais cette langue contient en parties à peu près égales des éléments de ces deux espèces. Cela tient, comme nous le verrons, à ce que la langue allemande est née dans les chancelleries et qu'elle ne représente pas un seul et même dialecte passé de l'état populaire à l'état littéraire.

Le principe fort simple du haut-allemand rigoureux est celui d'un nouveau renforcement.

Nous avons vu que les « gh, dh, bh » de la langue commune indo-européenne étaient devenus en gothique, en bas-allemand et dans les langues scandinaves « g, d, b » : le haut-allemand renforce ces derniers et en fait des *k*, *t*, *p*. Les « g, d, b » organiques étaient devenus « k, t, p » dans les langues germaniques du premier degré ; le haut-allemand les renforce à nouveau : il change « k » en *h* (qu'il écrit également *hh* ou *ch*), « p » en *f* (qu'il écrit également *pf* ou *ph*). Quant au « t », au lieu d'en faire un « th » sifflant, il le change en « ts », sous la forme *z*. Les explosives organiques « k, t, p » étaient devenues « h, th, f » dans les langues germaniques du premier degré : le haut-allemand conserve la première

et la dernière de ces sifflantes, ne pouvant les renforcer, et il applique à la seconde, au « th » dur, le phénomène de la polarité ; cette troisième série de consonnes se présente donc en haut-allemand sous la forme de *h*, *d*, *f*.

C'est pour ce motif que l'allemand présente un *d* là où l'anglais offre un *th ;* il dit, par exemple, *der* « le », *dorn* « épine », *trei* « trois », *dünn* « ténu, mince », tandis que l'anglais dit *the thorn*, *three*, *thin*. Dans ce cas comme dans tous les autres, l'anglais est ainsi plus pur d'un degré que ne l'est l'allemand ; *zæhmen* « dompter », *zæhre* « larme », *zu* « à, vers », *zwei* « deux » sont moins purs, sous ce rapport, que les mots anglais *tame*, *tear*, *to*, *two*. On commet une grave erreur en répétant que l'anglais provient de l'allemand ; autant dire que le gothique, lui aussi, en dérive. Ce sont là des branches parallèles ; mais le phénomène d'un second renforcement de certaines consonnes donne à l'allemand un caractère incontestable d'infériorité.

Tous les dialectes du haut-allemand ont changé en *t*, *z*, *d* les « d, th, t » des idiomes germaniques du premier degré. En cela ils ont tous été du « rigoureux » haut-allemand. Mais il n'en a pas été de même pour les deux autres ordres de consonnes. Une partie seulement des idiomes allemands changèrent les « k, g » du premier fonds germanique en *h*, *k*, et les « p, b » en *f p ;* c'est-à-dire qu'une partie seulement de ces dialectes réalisèrent dans toute sa rigueur le second renforcement. Tandis que le gothique, par exemple, dit *brinnan* « brûler », certains dialectes du haut-allemand disent *prinnan :* c'est là le haut-allemand rigoureux ; mais d'autres dialectes n'ont point renforcé le *b* et l'allemand littéraire actuel dit *brennen*. Tandis que le gothique dit *galeiks* « pareil, semblable », l'ancien haut-allemand rigoureux dit *kilih*, mais l'allemand littéraire dit *gleich*. Tandis que le gothique dit *kunnan* « connaître », le haut-alle-

mand rigoureux dit *chunnan* (avec ch=h) et l'allemand littéraire *kennen*. Mais, répétons-le, lorsqu'il s'est agi de la série des consonnes dentales, l'évolution s'est opérée dans tous les dialectes.

L'ancien haut-allemand reçoit parfois le nom de « tudesque ».

Il comprend trois grands dialectes, trois dialectes principaux, qui se subdivisent eux-mêmes en un assez grand nombre de dialectes moins importants. Ce sont le dialecte franc, le dialecte alaman-souabe, le dialecte austro-bavarois. Leurs monuments littéraires vont du septième siècle de notre ère jusqu'à la fin du onzième.

La grande caractéristique de ces dialectes réside dans le fait qu'ils conservent encore les anciennes voyelles dans les désinences : *nimu* « je prends », *nimit* « il prend », *nëmat* « vous prenez ». Nous verrons qu'à partir du douzième siècle ces différentes voyelles se sont changées en *e* ou bien sont tombées. Le vieux haut-allemand n'eut pour ainsi dire pas de littérature nationale ; il possède un certain nombre de traductions d'écrits religieux et des poésies chrétiennes, mais il n'a rien de véritablement germanique.

Avec le douzième siècle commence le moyen haut-allemand. La littérature revient aux anciennes traditions, aux anciennes fables que le vieux haut-allemand avait négligées, mais elle ne les envisage plus qu'à travers les idées et les conceptions chrétiennes. Cette période dure environ quatre cents ans. C'est l'âge des célèbres poètes « minnesænger », de Walther von der Vogelweide, de Wolfram von Eschenbach, de Nîthart, de Heinrich von Morungen, du Tanhûser.

La grande caractéristique de la langue de cette période est le changement en *e* de la voyelle des syllabes terminales : le vieux haut-allemand, le tudesque, dit *gibu* « je donne », le moyen haut-allemand dit *gibe*. Les diffé-

rents dialectes de l'ancien haut-allemand s'accommodèrent à cette nouvelle loi et continuèrent à garder chacun leur individualité et leur caractère particulier. Il se forma toutefois une sorte de langue littéraire, une langue des cours, tirée du dialecte souabe (1) ; pareille chose ne s'était point produite dans la période précédente.

Deux faits bien frappants, ajoute Schleicher, distinguent le moyen haut-allemand de l'allemand moderne. Dans le premier les syllabes radicales sont tantôt longues, tantôt brèves ; dans le second, la syllabe radicale est toujours longue ; c'est elle, comme l'on sait, qui porte l'accent : l'accentuation, en allemand moderne, détermine donc la longueur de la syllabe qu'elle affecte, c'est-à-dire de la syllabe radicale. Le second fait est celui-ci : « Dans le vieux haut-allemand nous n'avions jamais sous les yeux que le dialecte de celui qui tenait la plume ; il n'existait pas de langue littéraire d'un emploi plus général et dominant les différents dialectes. Durant la période du moyen haut-allemand il se forme une langue plus générale, la langue des cours. L'allemand moderne est encore moins un dialecte particulier que ne l'était la langue des cours du moyen haut-allemand ; ce n'était point la langue de telle contrée, elle n'était parlée par aucune population. Telle est la cause du manque de naturel que présente la langue allemande ; dans sa phonétique, dans ses formes elle est souvent monstrueuse. Mais aussi elle puise dans ce même fait de n'être point un idiome spécial, de n'appartenir en propre à aucune population particulière, la faculté de servir de lien d'union aux différentes branches germaniques... » (Op. cit., ibid.).

En remontant de nos jours jusqu'au temps de Luther, on peut suivre pas à pas la langue allemande. Sans

(1) Schleicher. *Die deutsche sprache*. Deux. édit., p. 103 et suiv. Stuttgard, 1869.

doute, durant cette période de plus de trois siècles, elle a subi bien des modifications, mais en réalité c'est toujours la même langue, c'est une seule et même langue. Au seizième siècle nous la voyons naître dans les chancelleries, nous voyons les actes diplomatiques emprunter arbitrairement aux différents idiomes populaires. Grâce à l'influence des actes officiels, grâce surtout à la propagande luthérienne, il se fait jour peu à peu ; il pénètre dans l'église, dans l'école, dans les tribunaux ; les dialectes populaires cèdent peu à peu devant lui et ne se défendent bientôt plus que dans les campagnes.

Il faut reconnaître d'ailleurs que la bizarre orthographe dont on l'affubla n'était point faite pour hâter sa propagation littéraire. Rien de plus arbitraire que cette orthographe. Parfois, pour allonger les voyelles, on les fait suivre d'un *h* qui ne répond absolument à rien dans le passé du mot que l'on défigure ainsi ; parfois, également pour indiquer qu'une voyelle est longue, on la redouble ; et comme, parfois encore, la voyelle est longue sans que sa quantité de longue soit figurée par un signe graphique quelconque, il arrive qu'un *a* long peut être rendu de trois façons différentes : simplement par *a*, par *ah*, par *aa*. C'est le cas des mots « zwar, wahr, haar ». Souvent, là où il faudrait un *i* pur et simple, on écrit *ie ;* souvent aussi, lorsque l'étymologie historique demanderait que l'on écrivît *ie*, l'on n'écrit que *i*. Souvent enfin, ce qui est tout aussi bizarre, on remplace les *t* par des *th*. Des tentatives sérieuses, mais incomplètes, ont été faites pour arriver à une réforme.

§ 7. Branche slave.

Les langues slaves ont occupé au moyen âge, durant les septième, huitième et neuvième siècles, de vastes régions de l'Europe centrale où l'allemand seul est connu aujourd'hui : la Poméranie, le Mecklembourg, le Brandebourg, la Saxe, la Bohême occidentale, la Basse-Autriche, la plus grande partie de la Haute-Autriche, la Styrie du nord et la Carinthie septentrionale. On parlait des idiomes slaves sur les lieux qu'occupent à présent Kiel, Lubeck, Magdebourg, Halle, Leipzig, Baireuth, Linz, Salzbourg, Gratz et Vienne.

On distingue ordinairement dans les langues slaves deux groupes principaux. Nous verrons tout à l'heure comment ils sont composés, et comment on a cherché à classer entre elles les différentes langues slaves ; mais il nous faut, auparavant, aborder une autre question générale, la question de la vieille langue ecclésiastique slave.

Dès le septième siècle, les populations slaves avaient atteint leurs limites extrêmes vers l'occident : le christianisme les attaqua de l'est et du sud, de Constantinople et de Rome (1). C'est aux Bulgares, aux Serbes, aux Russes que s'adressa la propagande partie de Constantinople, dont les résultats furent précoces. Avec le christianisme s'introduisit la liturgie en langue slave.

L'apostolat des frères Constantin (Cyrille) et Méthode donna à ce mouvement l'impulsion décisive. Ce fut vers le milieu du neuvième siècle que Cyrille réforma à l'usage des Slaves de Bulgarie l'alphabet grec, traduisit les Evangiles, un certain nombre de pièces liturgiques et se rendit avec son frère chez les Slaves de Moravie.

(1) Schafarik. *Geschichte der südslavischen litteratur*, t. III. Prague, 1865.

Méthode, évêque de Moravie et de Pannonie, mourut, après son frère, en 885. L'évangile d'Ostromir, qui date de 1056, est le plus ancien manuscrit de la langue dont se servirent Cyrille et Méthode et que l'on appelle, en raison de son emploi dans les offices religieux, *slave ecclésiastique*, *esclavon liturgique*. On lui applique également, comme nous le verrons tout à l'heure, un certain nombre d'autres noms.

La modification de l'alphabet grec due à Cyrille prit le nom d'écriture « cyrillienne ou cyrillique » ; elle est encore en usage, sous une forme très peu différente, chez les Russes, les Bulgares et les Serbes. Les Roumains, peuple de langue latine, avaient, eux aussi adopté cet alphabet, qu'ils ont heureusement rejeté aujourd'hui, pour en revenir aux caractères latins ; ils n'ont eu besoin que de leur ajouter un certain nombre de signes plus ou moins conventionnels.

Un jour viendra, il faut l'espérer, où la littérature russe fera à son tour le sacrifice de son alphabet traditionnel. Sans préjuger des circonstances qui pourront amener ce grand et fécond événement, on peut penser qu'elles ne se feront pas indéfiniment attendre ; la civilisation des deux parties de l'Europe trouvera dans cette réforme un accroissement considérable.

On se servit également, chez les Slaves du rite latin, d'un autre alphabet, dit « glagolitique ». L'origine de ce dernier est encore obscure ; quelques auteurs ont même prétendu qu'il était le plus ancien, mais l'opinion vraisemblable admise aujourd'hui communément est que le glagolitique n'est qu'une déformation du cyrillien : on prétend qu'il date de la fin du onzième siècle, et doit son origine au désir des Slaves du sud-ouest de sauver, grâce à l'emploi de signes incompréhensibles, leur liturgie qu'un concile avait prohibée. Quoi qu'il en soit de cette explication, il nous semble à peu près démontré

que l'alphabet glagolitique n'a point d'autre origine que l'écriture cyrillienne.

Il est impossible de répondre d'une façon précise à la question de savoir quelles étaient, au neuvième siècle, les limites géographiques du slave ecclésiastique. Les auteurs qui ont cherché à éclaircir ce point très obscur ne sont point arrivés aux mêmes conclusions. Selon les uns le slave ecclésiastique aurait été parlé dans le sud-ouest de la Russie actuelle, selon d'autres en Moravie, selon d'autres encore dans les régions de la Carinthie, de la Croatie, de la Slavonie, de la Serbie actuelle ; quelques-uns pensent qu'il s'étendait sur tout le territoire compris entre le Pont-Euxin et la mer Adriatique.

D'après Dobrovsky, dont l'opinion sera toujours d'un grand poids dans les questions de philologie slave, le slave ecclésiastique aurait été parlé en Serbie, en Bulgarie et en Macédoine. Sa frontière du nord aurait été le Danube ; au sud il se serait étendu jusqu'à Thessalonique.

Pour Schafarik (1), le vieux bulgare s'étendait, avant l'arrivée des Magyars, non seulement au sud du Danube, mais encore au nord de ce fleuve, sur la Valachie actuelle, sur le territoire des Saxons en Transylvanie, sur la Hongrie orientale. Plus tard il lui donna le nom d'ancien slave (en tchèque *staroslovansky*). F. Miklosich, qui l'appelle ancien slovène (*lingua palæoslovenica*), opine pour la région de la Dacie et du territoire hongrois situé sur les deux rives du Danube. Cette langue n'aurait donc pas été parlée dans la presqu'île des Balkans (2). Ajoutons que cette opinion est également celle de Danitchitch (3).

Le slave ecclésiastique a disparu entièrement en tant

(1) *Slovansky národopis*, p. 33. Prague, 1842.
(2) *Altslovenische formenlehre*. Introd. Vienne, 1874.
(3) *Dioba slovenskih jezika*. Belgrade, 1874.

que langue parlée, mais il a persisté, avons-nous dit, dans la liturgie. Ce n'est pas toutefois sans s'être quelque peu modifié, sans avoir subi, notamment, l'influence des idiomes vivants au milieu desquels on l'employait comme langue morte. Ces modifications sont relevées et connues ; de là deux formes du slave ecclésiastique : l'une ancienne, l'autre plus moderne. C'est de la première, bien entendu, que les linguistes ont à se servir si souvent dans l'étude des langues slaves, encore qu'il ne faille point la tenir (ainsi que nous le dirons plus loin) pour la mère de tous ces idiomes.

Les langues slaves vivantes sont le *russe*, le *ruthène*, le *polonais*, le *tchèque* et le *slovaque*, les deux dialectes du *sorbe* ou *sorabe*, le *bulgare*, le *serbo-croate* et le *slovène*.

Les limites de la langue *russe* vers le nord et vers l'est sont assez difficiles à déterminer. Elle s'y rencontre, en effet, avec les nombreuses langues ouralo-altaïques (samoyède, zyriénien, vogoul, etc.) qu'elle pénètre peu à peu. Du côté de la Baltique elle confine à peine au littoral, qu'occupent le suomi et l'esthonien, idiomes finnois, le suédois (Helsingfors) et le lette (Riga, Mitau) ; un peu plus au sud elle est limitrophe du lithuanien. De Grodno jusqu'à une centaine de lieues vers le sud, à peu près en ligne directe, elle a pour limite occidentale le polonais. Au sud, enfin, elle se rencontre avec le ruthène dont nous parlerons tout à l'heure.

Dans ces limites, nous comprenons, d'ailleurs, le dialecte dit « russe blanc », parlé par près de trois millions d'individus (au nord du ruthène, à l'ouest du russe, à l'est du lithuanien et du polonais), à Vitebsk, Minsk, Mohilev, et dont la littérature est fort peu importante.

Le grand russe, ou russe simplement dit, n'est pas tout à fait le même dans sa langue littéraire et dans sa langue vulgaire ; la première a fait des emprunts sen-

sibles à la langue slave ecclésiastique. Les plus anciens monuments du russe — que l'on peut suivre jusqu'au onzième siècle — sont des contes et des épopées. La langue se régularise tout à fait durant le dix-huitième siècle, grâce en partie au célèbre érudit et littérateur Lomonosov (1711-1766), et elle donne, depuis cette époque, des témoignages d'une originalité et d'une vitalité littéraire que l'on ne peut apprécier que trop rarement.

La grammaire du russe n'est malheureusement pas sans offrir d'assez grandes difficultés à qui ne connaît que les langues novo-latines ou les langues germaniques. Son matériel phonétique est assez complexe ; la prononciation des voyelles n'est pas toujours la même : ainsi *a*, dans les syllabes non accentuées, prend quelque peu le son de *e ; e* se prononce parfois ouvert, parfois fermé ; *o* se prononce *a* dans les syllabes inaccentuées : ainsi, dans le mot *kolokol* « cloche », l'accent étant sur la première syllabe, le premier *o* seul garde sa valeur et les deux autres se prononcent *a :* « kolakal ». L'accentuation russe, comme celle de quelques autres langues slaves, est d'ailleurs d'une difficulté considérable ; cette accentuation est bien connue, assurément, mais ses lois sont fort loin d'être toutes fixées. La déclinaison du russe est à peu près la même que celle des autres idiomes slaves, et l'on ne peut guère y signaler que l'usage des lois phonétiques plus ou moins spéciales à cette langue. Il se distingue dans sa conjugaison par la perte complète de deux des anciens temps, l'aoriste et l'imparfait (que le ruthène a également perdus, que le serbe et le bulgare ont conservés, dont les plus anciens monuments polonais montrent encore des traces et que possédait la vieille langue tchèque). Il les remplace par un participe : *on dal* « il a donné » (masculin), au féminin *dala*, au neutre *dalo*, au pluriel *dali* pour les trois genres ; cette formation périphrastique a en quelque sorte le sens

de « je suis ayant donné, nous sommes ayant donné ».

Le *ruthène*, également appelé *rusniaque* et *petit russe*, n'est pas un dialecte du russe, bien qu'il s'en rapproche plus que de toute autre langue slave. Il occupe, approximativement, un cinquième du territoire de la Russie d'Europe. A l'ouest il confine au polonais, au nord-ouest il confine au russe blanc, au nord-est et à l'est il confine au russe proprement dit. Il est parlé également à l'est de la mer d'Azov. En Autriche il s'étend sur la plus grande partie de la Galicie et forme la bande nord-orientale de la Hongrie, au-dessus du magyar et du roumain. Les Ruthènes de Russie sont au nombre d'environ onze millions et demi, y compris les Cosaques ; ceux d'Autriche-Hongrie sont évalués à plus de trois millions cinquante mille, ce qui donne un total de plus de quatorze millions et demi d'individus parlant le petit russe.

Leur littérature, comme celle des Slaves du sud et aussi comme celle des Russes, est avant tout une littérature populaire et traditionnelle. On a publié, depuis une cinquantaine d'années, sous les noms de chants populaires de l'Ukraine, chants populaires de la Russie méridionale, de la Galicie, de la Volhynie, un grand nombre de morceaux en langue ruthène.

Pour ne pas être très différent du russe, le ruthène s'en distingue cependant d'une façon fort nette. Il ne « mouille » point, par exemple, toutes les consonnes que le russe peut mouiller, entre autres les labiales *p*, *b*, *v*, *m* ; il change plus facilement que le russe les « k » et les « g » de l'antiquité en *tch* et *j* (le « j » français) ; il a une accentuation souvent différente ; il a perdu le participe présent passif, que le russe a conservé ; il possède une forme d'infinitif à sens diminutif. Ces particularités, ainsi que bien d'autres faits plus ou moins notables, ont suffi à le faire regarder comme un idiome indépendant et bien caractérisé.

Le *polonais* comprend un certain nombre de dialectes, dont l'ensemble couvre un assez vaste territoire réparti entre la Russie, la Prusse et l'Autriche. La limite orientale est assez connue ; elle va de Grodno à Jaroslav, en longeant une partie du Boug. Celle de l'ouest est moins précise ; l'allemand envahit chaque jour, de ce côté, le territoire de langue polonaise et en occupe toutes les localités un peu importantes ; en Autriche la Galicie occidentale est polonaise : cette région est beaucoup moins grande que la partie orientale du même pays, occupée, ainsi que nous l'avons dit plus haut, par les Ruthènes. Le polonais occupe, en somme, une espèce de parallélogramme — très irrégulier sur son flanc occidental — dont la hauteur est en moyenne de cinq cents kilomètres, et dont la plus grande largeur (par la ligne Posen-Varsovie) n'est pas tout à fait aussi considérable. La langue allemande a beaucoup gagné sur le polonais ; toute la région occidentale de cette langue, même sur le territoire russe, est parsemée de petits ilôts où l'on ne parle qu'allemand : il s'en rencontre quelques-uns presqu'aux portes de Varsovie, et la Galicie n'est point préservée de cette invasion, due principalement au progrès de la population israélite.

Le nombre des Polonais de Russie est évalué à 4 millions 700.000, celui des polonais prussiens à 2.450.000, celui des Polonais d'Autriche et de Hongrie à 2.465.000, ce qui donne un total approximatif de 9.615.000 individus parlant polonais.

La phonétique du polonais est assez compliquée, et l'alphabet qui lui est appliqué peut passer pour un des plus défectueux ; c'est ainsi que le son « tch », au lieu d'y être rendu par un seul signe — comme en tchèque et en croate (*c*) — y est exprimé par le groupe *cz* ; au lieu du *s* tchèque et croate, qui a la valeur de notre « ch » et du « sh » anglais, le polonais possède le groupe *sz* ;

pour le *v* croate (notre « v »), il met *w* à la façon allemande. Ce ne sont pas là, d'ailleurs, les seules incommodités de sa transcription, et si les tentatives de réforme qu'on semble vouloir lui appliquer en ce moment arrivent à bon terme, il y aura lieu de s'en féliciter grandement à tous les points de vue.

Outre les voyelles *a*, *e*, *i*, *o*, *u* (« ou » français), *y*, son étranglé qui se rapproche de notre « u », outre un *é* très rapproché du son « i », un *ó* analogue à notre « ou », le polonais possède deux voyelles nasales dont la valeur approximative est celle de « an » et « in » français (chant, fin). En certains cas, notamment à la fin des mots, elles prennent la valeur de *o* et de *e*. En somme, elles répondent à deux voyelles nasales de l'ancien slave ecclésiastique, qui, selon toute vraisemblance, correspondaient à nos nasales françaises « on » et « in ».

Les variations auxquelles se trouvent soumises les consonnes de la langue polonaise,selon la rencontre qu'elles font de telle ou telle autre consonne, sont assez importantes. C'est ainsi que les sifflantes subissent des permutations capables de rendre souvent très obscure l'origine des mots. Quant à l'accentuation, elle est fort simple : elle porte toujours sur l'avant-dernière syllabe, sauf dans les mots empruntés aux langues étrangères.

Nous avons vu qu'en russe et en ruthène l'accent pouvait tomber sur toute syllabe, indépendamment de la place de cette syllabe dans le mot : nous verrons qu'il en est de même en slovène et en croato-serbe ; en tchèque et en sorabe il affecte la première syllabe. Le polonais, sous ce rapport, est donc bien caractérisé.

La littérature polonaise est non seulement importante, elle est encore originale. Elle commence à la fin du dixième siècle et compte une foule de chroniqueurs et de poètes à partir du douzième siècle. Elle est encore aujourd'hui des plus importantes. Une bibliographie, pu-

bliée à l'occasion de la dernière exposition de Vienne, porte à plus de trois mille le nombre des ouvrages imprimés en polonais ou publiés par des Polonais en langue étrangère pendant la seule année 1871.

Les limites actuelles du *tchèque* et du *slovaque*, qui lui est intimement allié, sont difficiles à décrire. La région qu'ils occupent (toute la Bohème, moins une lisière de l'ouest et du nord, la plus grande partie de la Moravie et le pays situé au sud du territoire de langue polonaise) s'étend de Pilsen aux Carpathes sur une longueur d'environ cent cinquante lieues. On compte environ 7.300.000 Tchèques, Moraves et Slovaques (*Archiv f. Slavische Philol.*, tome II, p. 407).

Depuis les premiers monuments que l'on en possède et qui datent du huitième siècle, la langue tchèque a subi d'importantes modifications ; c'est un fait que nous devons attribuer aux mouvements politiques si considérables dont la Bohême a été le théâtre.

Nous ne faisons pas allusion ici à la simple différence d'ortographe, résultant de ce que dans les premiers documents tchèques on se servit des caractères latins tels quels, sans les modifier, à l'occasion, par des signes accessoires indispensables, des signes diacritiques, mais nous parlons des formes elles-mêmes de la langue. La réforme orthographique du tchèque, commencée il y a plusieurs siècles, s'est complétée en 1830 par l'abandon des caractères gothiques, et s'est définitivement achevée, il y a une trentaine d'années, par la substitution du *v* latin au *w* des Polonais et des Allemands. Cette réforme, qui s'imposait impérieusement, fut d'une importance capitale pour la langue elle-même, pour son développement, pour sa propagation. Rien n'était moins fixé que l'ancienne écriture tchèque ; un seul et même son était souvent transcrit de trois, quatre, cinq et six manières différentes. Ainsi *s* était indifféremment rendu par « z,

s, sz, szs, zz, ss », *k* était transcrit par « c, k, q, ch, k, ck, », et ainsi de suite ; un même caractère latin, par contre, rendait souvent trois ou quatre sons tout à fait différents. L'on conçoit combien il est difficile, avec un pareil système, ou plutôt avec une pareille absence de système, de rétablir exactement les textes tchèques.

Les voyelles tchèques *a*, *e*, *i*, *o*, *u* (prononcez « ou »), *y* (qui ordinairement se prononce *i*) ont toutes leurs longues que l'on distingue dans l'orthographe actuelle par le signe « minute » : *â*, *é*, etc. Une autre voyelle tchèque, *e*, possède également sa longue, mais il n'y a point pour celle-ci de signe distinctif. Cette voyelle se prononce « yé ». Le tchèque possède encore un *r* voyelle et un *l* voyelle, toujours brefs dans le dialecte ordinaire, mais qui, en slovaque, peuvent être longs ; quant aux voyelles nasales du polonais, elles lui sont inconnues et l'on n'a pu les retrouver même dans les plus anciens textes. Les voyelles tchèques sont assez variables et subissent d'une façon très sensible l'influence qu'exerce sur elles le voisinage d'un *j* (prononcez « y »), qui change, par exemple, en *e* et en *i* les *a* et les *e* dont il est suivi, et en *e* les *a* dont il est précédé.

Le système des consonnes tchèques est très riche : il possède des dentales mouillées ; un *r* particulier, correspondant au *rz* polonais, ayant la valeur de « rj » français, et que l'on rend par le caractère *r* ; des sifflantes très sensibles au contact de certains autres sons.

Nous avons dit plus haut que le tchèque accentuait toujours la première syllabe de chaque mot.

Notons enfin que la conjugaison de l'ancien tchèque était d'une grande conservation ; la langue moderne a perdu (comme presque tous les autres idiomes slaves) l'imparfait et l'aoriste anciens.

La littérature tchèque remonte, nous l'avons dit, au huitième siècle. Ses premiers monuments sont les célè-

bres manuscrits de Královdor (Kœniginhof) et de Zelenohora (Grünberg), découverts en 1817, et dont l'authenticité est avérée. Ils remontent à la période de transition entre le paganisme et le christianisme et sont aussi importants pour la linguistique que pour l'étude des anciens mythes religieux de la Bohême. On possède également plusieurs fragments datant du dixième siècle. Jusqu'à l'époque de la guerre des Hussites, la Bohême, qui la première avait donné le signal de l'émancipation religieuse, posséda la plus importante de toutes les littératures slaves. Quand elle fut tombée sous la domination allemande, sa langue nationale fut sévèrement proscrite, et quiconque essaya de la remettre en honneur devint la victime des Jésuites. Ce n'est que depuis la fin du siècle dernier que les lettres bohêmes ont reçu une vie nouvelle.

Le *sorbe* ou *sorabe*, également appelé *vinde* ou *serbe de Lusace*, comprend deux dialectes distincts, le bas-sorabe et le haut-sorabe. L'ensemble de cet idiome n'occupe plus, aujourd'hui, qu'un territoire d'environ vingt-cinq lieues de hauteur, traversé par la Sprée, sur dix ou douze de largeur : les deux tiers de la région sont situés en Prusse, le tiers méridional en Saxe, et les localités les plus importantes (Kottbus, Bautzen) sont envahies par l'allemand. Un espace d'à peu près douze lieues sépare la frontière sorabe méridionale de la frontière tchèque septentrionale. Vers le milieu du seizième siècle la contrée où se parlait le serbe de Lusace était deux fois plus considérable qu'elle ne l'est aujourd'hui. C'est par le nord, particulièrement, par l'ouest et par l'est que la langue allemande a empiété peu à peu sur ce domaine, qui ne peut guère revendiquer à l'heure actuelle plus de 130.000 habitants de langue slave.

Le plus ancien document imprimé en langue vinde est un livre de prières catholiques publié en 1512. Le dix-

septième siècle compte un certain nombre d'écrivains en langue sorabe, mais au commencement du dix-neuvième cette littérature était presque entièrement abandonnée. On entreprit plus tard de la remettre en honneur, et en 1845 fut créée une société qui est devenue le centre de la vie littéraire du pays.

La langue serbe, ou croate, ou mieux *serbo-croate*, avec ses deux grands centres intellectuels, Belgrade et Zagreb (Agram), occupe une place considérable, non seulement parmi les idiomes sud-slaves, mais encore parmi les langues slaves en général. Cette place lui est légitimement dévolue au triple point de vue de la linguistique, de l'histoire, de la géographie. Les pays sur lesquels elle s'étend sont la principauté de Serbie, la Bosnie, l'Herzégovine, le Monténégro, une partie de la Hongrie méridionale (Zombor), la Slavonie, la Croatie, la presque totalité de l'Istrie, la Dalmatie ; c'est une région comprenant près de 6 millions d'habitants.

Sur un territoire aussi étendu les sous-dialectes sont assez nombreux, mais on peut dire qu'il existe trois dialectes principaux : celui de l'ouest, moins littéraire que les deux autres ; celui du sud, qui fleurit surtout en Dalmatie ; celui de l'est, parlé dans la principauté serbe et dans la Hongrie méridionale sur les rives du Danube. La caractéristique principale de ces trois dialectes est la prononciation différente d'une voyelle qui dans l'antiquité se prononçait très certainement *é* : à Belgrade, dans la Hongrie du sud, en Sirmie, on lui a conservé cette valeur ; le dialecte de l'ouest la change en *i* ; celui du sud en *je* ou *ije* (prononcez « yé, iyé »). Mais que l'on dise *vera* ou *vijera* « la croyance » — prononcez « véra, viyéra », — *reka*, *rika*, ou *rijeka* « la rivière », l'on sera compris sans peine de l'Adriatique jusqu'à la frontière roumaine.

Le malheur de la langue croato-serbe est de posséder

un double alphabet : à l'est l'alphabet cyrillien, à l'ouest l'alphabet latin complété à l'aide de certains signes accessoires. Cette division regrettable est la conséquence de l'ancienne scission religieuse ; on ne saurait trop la déplorer. Elle retardera longtemps encore les rapprochements de toute espèce que la civilisation européenne aurait tant d'intérêt à voir s'opérer entre les Serbes de Turquie et le royaume tri-unitaire dalmato-croato-slavon.

Ce n'est point qu'un pas considérable n'ait été fait dans cette voie, au commencement du siècle, par l'espèce d'unification et de codification que le célèbre Vouk Stéphanovitch Karadjitch introduisit dans le langage des Serbes de la Principauté et de la Hongrie méridionale.

Lorsque Vouk entreprit l'œuvre considérable qu'il lui fut donné de mener à si bonne fin, on pouvait dire que la langue serbe n'était pas encore définie. Presque tous les lettrés regardaient comme l'idiome national une langue assez factice formée d'éléments de l'ancien slave liturgique mélangés avec des éléments de la langue réellement vivante et populaire. Cette dernière ne passait à leurs yeux que pour un simple patois. Vouk proposa d'adopter telle quelle la langue nationale et de réformer radicalement son orthographe. Ce fut une lutte d'un demi-siècle. Il en sortit vainqueur, grâce à sa parfaite connaissance de sa langue croato-serbe, grâce à la précision et à la méthode de ses travaux.

Le fond de la littérature serbo-croate est le chant populaire, le chant national ; *pjesma*, *pisma*, *pesma*. Un grand nombre de ces morceaux ont été recueillis et publiés ; beaucoup d'entre eux sont sans doute très anciens, et la forme même sous laquelle ils se présentent laisse voir combien la langue s'est peu modifiée depuis des siècles.

Ce n'est point dans sa grammaire qu'elle a souffert,

c'est dans son lexique, au milieu duquel, par exemple, le dialecte de l'est a admis un nombre par trop considérable de mots turcs. Ajoutez l'invasion dans la langue actuelle scientifique et littéraire de termes tirés de l'allemand ou du français.

La Serbie et les pays slaves de rite oriental ont eu leur mouvement littéraire particulier. Pour être peu connu, il n'a pas été sans importance. Il date au moins du commencement du treizième siècle, bien que les documents qui remontent à cette époque soient en eux-mêmes d'une valeur peu considérable. Avant le treizième siècle, tout au plus peut-être avant le douzième, l'on ne possède de l'idiome serbe que des séries de mots et de noms propres tirés presque tous d'auteurs grecs ou latins.

Les monuments écrits des pays croato-serbes occidentaux remontent au douzième siècle, mais c'est au seizième seulement qu'ils se multiplient et que se développe la littérature dite ragusaine, du nom de Raguse qui en fut le centre. La littérature spécialement croate, qui occupe une place importante dans le domaine de la critique historique et de l'étude du langage, n'apparaît guère qu'à la fin de ce siècle.

L'étude particulière de la langue croato-serbe est d'une importance capitale dans l'étude générale des langues slaves, et l'on peut dire qu'elle vient immédiatement après celle du slave ecclésiastique. De tous les idiomes de cette famille, c'est en effet le croato-serbe (avec le slovène) qui a eu le moins à souffrir dans sa phonétique, et ce sont précisément les questions de phonétique, ainsi que nous l'avons vu, qui forment la base de toute étude linguistique. La grammaire comparée des langues slaves de F. Miklosich, ouvrage fondamental pour l'étude des idiomes de ce groupe (1), fournit à chaque instant la

(1) *Vergleichende grammatik der slavischen sprachen.* Vienne, 1852.

preuve éclatante d'énorme importance du croato-serbe, et la lecture des excellents travaux de Danitchitch, Jagitch, Novakovitch, dont la traduction en français rendrait un grand service, enlèverait sans peine les derniers doutes qu'il soit possible d'avoir à ce sujet.

Le matériel phonique de la langue serbe n'est pas compliqué. Il comprend six voyelles : *a*, *e*, *i*, *o*, *u* (« ou » français) et *r*. Son système de consonnes est fort simple ; nous les possédons toutes en français, sauf les deux palatales mouillées que l'on transcrit l'une par le signe *c*, l'autre par le groupe *gj* (nous préférerions le signe *j'*). La première de ces consonnes a la valeur d'un « t » suivi du « ch » allemand de « ich, brechen », la seconde équivaut à un « d » mouillé d'une façon analogue. L'accentuation du serbe est des plus difficiles pour un étranger : on compte pour l'ordinaire quatre espèces d'accents, bien qu'il faille, en réalité, les réduire à deux : l'un fort, l'autre faible, et chacun d'eux tantôt bref, tantôt long. Une grande supériorité du croato-serbe sur presque toutes les autres langues slaves, c'est qu'il a conservé les anciens aoristes et imparfaits : *bih* « je fus », *bijah* « j'étais », tout en possédant un passé formé à l'aide d'un participe : *sam bio*, *smo bili* « j'ai été, nous avons été ».

Le *slovène*, parlé par plus de douze cent mille individus dans la Carinthie et la Styrie méridionales, dans la Carniole et une partie du nord de l'Istrie, est intimement allié au croato-serbe et partage son importance sous le rapport linguistique. Sa littérature écrite remonte au milieu du seizième siècle ; elle ne fut pas sans valeur, mais les progrès du serbo-croate ne lui promettent sans doute plus un avenir bien brillant. Les livres protestants imprimés à Tubingue sont le monument le plus important de la littérature slovène du seizième siècle. Durant les deux siècles suivants, elle trouva des représentants éminents. Murko et Kopitar illustrent leur époque, mais

ce dernier a écrit ses livres en allemand. Son exemple a été suivi par son compatriote et élève F. Miklosich, dont les travaux, qui dominent aujourd'hui la science slave, peuvent par ce fait être revendiqués par la science germanique.

Le *bulgare* occupe la plus grande partie de la Turquie européenne : au nord il longe le Danube, de Vidin à Silistrie, et même quelque peu au delà ; à l'est il a pour frontière l'Albanie ; au sud, il n'est séparé des mers Egée et de Marmara que par les bandes littorales où l'on parle grec ou turc ; à l'est il approche souvent de la mer Noire et partage avec le turc la région de l'extrême nord-est de l'empire. L'on arrive aisément pour les Bulgares au chiffre de cinq millions cinq cent mille individus, si l'on tient compte de ceux qui habitent la Russie du sud-ouest, et la Bessarabie cédée à la Roumanie par le traité de Paris (1).

Le bulgare moderne est de toutes les langues slaves celle dont les formes se sont le moins bien conservées, Il présente cette particularité — qui lui est commune, du reste, avec le roumain et l'albanais — qu'il possède un article placé à la fin des noms. Son vocabulaire a grandement subi l'influence des langues voisines : du turc, du grec, de l'albanais, du roumain. En tous cas, malgré l'altération de ses formes, le bulgare moderne offre des restes des anciennes nasales slaves qui ont totalement disparu des autres idiomes méridionaux.

La littérature bulgare ne date que d'hier. Jusqu'au milieu du XIXe siècle, les rares écrivains originaires de Bulgarie se servaient du russe ou de l'ancienne langue liturgique imprégnée de russe. Dans ces derniers temps, nombre de jeunes Bulgares se sont instruits. Aujourd'hui ils possèdent des journaux et leur littérature s'enrichit d'an-

(1) JIRETCHEK. *Geschichte der Bulgaren* (trad. du tchèque), p. 578. Prague, 1876.

née en année. Les entraves apportées par les Turcs au développement des nationalités européennes de la Turquie forcent malheureusement les Bulgares à s'instruire à l'étranger et à y faire paraître leurs livres ; une société littéraire, qui joue aujourd'hui un rôle important, a été fondée à Braïla, en Roumanie.

Terminons enfin cette énumération en citant les anciens dialectes du *slave de l'Elbe*, connus sous le nom de *polabe*, dialectes éteints actuellement, et dont les rares monuments (sur lesquels la langue allemande a exercé une influence considérable) datent de la fin du dix-septième siècle et du commencement du dix-huitième.

Nous avons dit plus haut quelle était l'importance considérable du *slave* ecclésiastique pour l'étude des autres langues de la même famille. Il ne faudrait pas s'attendre, cependant, à trouver dans la grammaire de cet idiome une image très fidèle de l'ancienne langue commune indo-européenne.

Sa phonétique est sujette à des modifications bien autrement graves que ne l'est celle du lithuanien ou celle du grec. A la vérité, son système vocalique n'est pas des plus compliqués — bien que la nasalisation fréquente de certains sons y soit une preuve incontestable de décadence et que les voyelles terminales des mots s'y trouvent gravement atteintes (par des principes d'ailleurs très fixes) ; mais ses consonnes sont soumises à des lois d'attraction et d'assimilation très nombreuses et très délicates. Ce n'est pas une des moindres difficultés que l'on rencontre dans l'étude des langues slaves. A une série de lois phonétiques assez complexes, ajoutez la multiplicité des consonnes. On peut dire que c'est spécialement dans les langues slaves qu'il importe avant tout de se rendre un compte exact des éléments phoniques et des règles qui régissent leur rencontre. Sans doute la conjugaison est relativement simple, mais la déclinaison

s'est trop souvent écartée des anciennes formules de l'indo-européen commun et la complication des lois phoniques qui se présentent souvent dans la rencontre du thème et des désinences ajoute à cette difficulté des difficultés nouvelles.

Jetons un coup d'œil rapide sur la grammaire de cet antique idiome.

Le slave ecclésiastique possède les voyelles *a*, *e*, *i*, *o*, *u* (« ou » français), *y* (vraisemblablement « u » français), un « é » fermé (transcrit *e* ou *é*) qui prit parfois la valeur de la syllabe française « ya » ; de plus un *i* et un *u* (« ou » français) très peu sonore, en quelque sorte étouffés (transcrits *i* et *u*) ; enfin les deux nasales transcrites *a* et *e*, qui équivalaient vraisemblablement, quant à leur prononciation, la première à « on » (de « bon, son »), la seconde à « in » (de « vin, cinq »).

Les diphthongues de la langue commune indo-européenne ne subsistent plus en slave liturgique ; elles s'y sont condensées en de simples voyelles. Le hiatus est volontiers évité, et cela, pour l'ordinaire, par l'intercalation d'un *j* (la demi-voyelle française « y ») ou d'un *v* purement euphonique. Ces *j* et *v* euphoniques se placent également en tête des mots qui commencent par une voyelle. Tandis, par exemple, que, fidèles au type commun, le grec dit *este* « vous êtes », le latin *estis*, le lithuanien *este*, etc., le slave liturgique dit *jeste*. C'est ce qu'on appelle en termes techniques la « pré-iotation », caractéristique remarquable de toutes les langues slaves : tchèques et serbe *jeste* (d'où *ste*).

Arrivons aux consonnes. Le slave liturgique (et tous les autres idiomes slaves ont agi de même) a transformé en explosives simples les explosives aspirées « gh, dh, bh » de l'indo-européen commun ; il les change en *g*, *d*, *b*. Par contre, il s'est formé un certain nombre de sifflantes inconnues à la langue commune indo-européenne,

ce sont nos « j, ch, z » français, et sous l'influence de lois phonétiques rigoureuses il a dû changer souvent en « tch » (que l'on transcrit par le signe *c*) des « k » primitifs.

L'ensemble des lois phonétiques auxquelles il se trouve soumis, a pris chez lui une extension considérable ; l'extension, même rapide, des lois d'assimilation dans les langues slaves — assimilation complète ou incomplète, assimilation d'une consonne avec la consonne précédente ou avec la consonne suivante — doit précéder toute autre question dans l'étude des langues slaves. Faute d'avoir une idée, au moins générale, de ces lois, on peut se créer les conceptions les plus fausses sur la formation des mots.

Le principe qui concerne la chute des consonnes à la fin des mots est également d'une grande importance : toute consonne terminale doit tomber en slave ecclésiastique. Tandis, par exemple, que le sanskrit dit *sûnus* « filius », *sûnum* « filium » (en allemand « sohn »), le slave ecclésiastique dit *synû* aux deux cas, laissant ainsi tomber soit la désinence *s* du nominatif, soit la désinence *m de* l'accusatif.

A côté de la déclinaison nominale ordinaire (substantifs, adjectifs, participes, noms de nombre et quelques pronoms) et de la déclinaison pronominale, le slave liturgique possède une déclinaison dite *composée*, particulière également au lithuanien et (avec un élément différent) aux langues germaniques. Cette déclinaison est composée des formes ordinaires de l'adjectif auxquelles s'ajoute le pronom *i* également décliné. En principe, les adjectifs admettent les deux déclinaisons, la déclinaison normale et la déclinaison composée ; l'emploi de l'une ou de l'autre est une question de syntaxe : décliné de la seconde façon, l'adjectif est dit *défini* et a le sens de l'adjectif grec ou allemand précédé de l'article. Tous les

idiomes slaves possèdent cette déclinaison composée ; le serbe, par exemple, dit : *rast visok* « un chêne élevé », *visoki rast* « le chêne élevé ».

Le slave ecclésiastique a conservé dans sa conjugaison les trois nombres de l'indo-européen commun : singulier, duel, pluriel ; le duel n'existe plus en croato-serbe, en bulgare, en ruthène, en russe. Des quatre temps simples de l'indo-européen commun, le slave liturgique a perdu le parfait redoublé (grec λέλοιπα « j'ai laissé »), l'imparfait ; mais il a conservé presque toutes les différentes formes du présent et l'aoriste. Il a conservé également les deux temps composés de la langue commune, aoriste et futur — au moins en partie. Par contre, il s'est forgé un imparfait assez composé.

De toutes les langues slaves actuellement vivantes, le serbo-croate et le slovène, son très intime allié, possèdent la phonétique la plus claire et la plus simple. Ce n'est pas à dire que les lois euphoniques si nombreuses qui concernent la rencontre des consonnes et dont nous venons de parler à propos du slave ecclésiastique, ne se présentent pas en serbo-croate. Elles y sont, au contraire, aussi exigeantes que dans tous les autres idiomes slaves, mais le matériel phonique lui-même est beaucoup moins compliqué dans cette langue que dans les autres idiomes de cette famille, et sa prononciation, en outre, n'offre aucune difficulté. Par contre, le polonais et le tchèque présentent, sous ce rapport, des obstacles sérieux. Quant au bulgare, les modifications qu'il a subies dans le cours des temps en ont fait la moins bien conservée de toutes les langues slaves.

La classification des langues slaves a donné lieu à des controverses importantes, et l'on peut dire que cette question n'est pas encore résolue.

On avait supposé tout d'abord que le slave ecclésiastique était la source commune de tous les idiomes de

cette famille ; de là les noms de « paléoslave », d' « ancien slave », dont on se sert parfois encore pour désigner cette langue. C'était une erreur grave. Il n'est personne aujourd'hui, parmi ceux qui s'occupent de grammaire slave, qui songe à soutenir encore cette opinion. Mais, après avoir écarté cette prétendue paternité du slave liturgique, fallait-il en venir à placer cet idiome sur le même pied que les langues de sa famille et supposer qu'ils étaient tous sortis directement d'une ancienne forme commune aujourd'hui perdue ?

On ne s'arrêta pas à cette hypothèse.

Dobrovsky et Schafarik divisèrent les langues slaves en deux branches principales : l'une occidentale, comprenant le polonais, le tchèque, le serbe de Lusace, l'ancien polabe ; l'autre sud-orientale, comprenant tous les autres dialectes. Schleicher commença par faire quelques objections contre cette division ; mais il finit par l'adopter, et nous pouvons résumer dans le tableau suivant son opinion à ce sujet :

- Slave commun.
  - Branche sud-orientale.
    - Sud-slave.....
      - Bulgare ancien et moderne.
      - Serbo-slovène.
        - Serbe.
        - Slovène.
    - Slave oriental
      - Grand russe.
      - Petit russe.
  - Branche occidentale.
    - Tchèque.
    - Polonais.
    - Sorbe.
    - Polabe.

On peut dire que Schleicher n'appuyait cette division que sur un seul fait : les *d* et les *t* placés devant un *n* ou un *l* tombent dans le premier groupe, tandis que dans le second groupe ils se conservent ; c'est ainsi, par exemple, que le tchèque *oradlo* « instrument de labour » est plus correct que le slave ecclésiastique *oralo* et le serbo-croate *oralo, ralo.*

G. Danitchitch n'admit point cette raison ; il démontra

que ces *d* et *t* tombent parfois aussi bien dans le tchèque ancien et moderne, ainsi qu'en polonais et en sorbe, et il fit voir également qu'en slave ecclésiastique et serbo-croate ils ne tombaient pas toujours.

F. Miklosich n'accepta pas davantage cette classification. Tandis que Schleicher regarde le slave ecclésiastique comme l'ancienne forme du bulgare actuel et lui donne le nom d'*ancien bulgare*, F. Miklosich pense que ce vieil idiome a pour représentants actuels, non seulement le bulgare, mais encore le slovène, et il l'appelle *ancien slovène*. Cette théorie fut vivement combattue par Schleicher, qui établit victorieusement, selon nous, par des arguments tirés de la phonétique, que le slovène actuel ne pouvait dériver de l'ancien slave ecclésiastique, et qu'il fallait, d'autre part, réunir en une seule branche le croate-serbe et le slovène ; c'était d'ailleurs l'idée de Schafarick (1).

C'est aussi en invoquant des raisons purement phonétiques que F. Danitchitch a établi récemment une très ingénieuse classification des langues slaves. Son mémoire, écrit en serbe, est malheureusement accessible à peu de lecteurs. Nous pouvons en résumer les conclusions dans le tableau que voici :

- Slave commun
  - *Polonais*, avec le dialecte *polabe*.
  - *Tchèque*, avec le dialecte *sorbe*.
  - Slave ecclésiastique.
    - *Ruthène*.
    - *Russe*.
    - *Bulgare*.
    - *Slovène*.
  - *Croato-serbe*.

On a proposé d'ailleurs plusieurs autres classifications, et il est vraisemblable que l'on en proposera encore de nouvelles. Aux deux tableaux qui précèdent nous

(1) Schleicher, *Ist das altkirchenslawische altslowenisch ?* Beiträge zur vergleichenden sprachforschung, t. I, p. 319.

pouvons joindre le suivant, auquel un certain nombre d'auteurs paraissent s'arrêter :

- Slave commun.
  - Branche du sud-est.
    - Russe.
      - Russe.
      - Ruthène.
      - Russe blanc.
    - Bulgare.
      - Slave liturgique.
      - Bulgare.
    - Serbo-slovène.
      - Croato-serbe.
      - Slovène.
  - Branche de l'ouest.
    - Tchèque et slovaque.
    - Polonais.
    - Serbe de Lusace ou Sorbe.
    - Polabe.

Ajoutons d'autre part que pour Johannes Schmidt c'est toute peine perdue que de vouloir dresser un semblable tableau. Toutes ces subdivisions sont purement théoriques, en fait elles n'ont jamais existé et les différents idiomes slaves ont procédé individuellement et peu à peu à leur propre formation (1). Nous reviendrons ci-dessous, et d'une façon plus générale, sur les prétendues subdivisions de la langue indo-européenne commune.

Quoi qu'il en soit à l'égard des langues slaves, et si la question, à vrai dire, nous semble encore obscure, deux points définitivement établis nous paraissent être l'antiquité des formes du serbo-croate et la grande détérioration du bulgare moderne.

Quant aux degrés de parenté plus moins étroits qui relient ces différents idiomes, quant aux formes communes plus ou moins intermédiaires qui auraient existé à une certaine époque, — par exemple un idiome commun tcheko-polono-sorbe, — nous n'en pouvons rien dire; nous n'en pouvons au moins rien assurer. Peut-être l'ave-

(1) *Zur geschichte des indogermanischen vocalismus*. Deuxième partie, p. 178. Weimar, 1875.

nir confirmera-t-il une part, sinon la totalité, de ce que l'on a écrit à ce sujet ; peut-être aussi en viendra-t-on un jour à ne plus voir dans tous les idiomes slaves qu'une série de collatéraux issus directement d'une source commune (sauf, selon toute vraisemblance, en ce qui concerne le bulgare moderne, qui proviendrait de l'ancien slave liturgique). Sans doute cela n'empêcherait pas le ruthène d'être moins dissemblable du russe qu'il ne l'est du slovène et du sorbe, cela n'empêcherait pas le polonais d'être moins dissemblable du tchèque qu'il ne l'est du bulgare ou du ruthène ; mais cela pourrait bien être fort exact.

En l'absence de documents historiques, il faut se montrer très réservé lorsqu'il s'agit de classifications de cette nature. C'est ce que nous pensons à l'égard des grandes divisions linguistiques, c'est ce que nous pensons également à l'égard des divisions plus particulières, entre autres celles des langues slaves.

§ 8. **Branche lettique.**

Sur la côte sud-est de la mer Baltique, dans les provinces russes de Courlande et de Covno, et dans l'extrême nord-est de la province allemande de Prusse orientale, il existe encore un petit groupe d'idiomes indo-européens, pressés à l'ouest par l'allemand, au sud par le polonais et le russe, à l'est par le russe également, au nord par une langue ouralo-altaïque, l'ehste, et qui est appelé à disparaître un jour ou l'autre devant le russe et devant l'allemand. C'est le groupe des langues lettiques. Jadis il était représenté par trois branches : le *vieux prussien*, le *lithuanien*, le *lette*. Il ne l'est plus aujourd'hui que par ces deux derniers ; le prussien a péri, il y a deux cents ans.

De toutes les langues indo-européennes, les langues

lettiques sont celles qui, en Europe, reflètent avec la plus grande fidélité l'ancien type commun indo-européen. Nous nous arrêterons sur le lithuanien plus que sur les deux autres ; c'est en effet l'idiome le plus important de ce groupe.

I. *Lithuanien.*

On compte en Allemagne cent cinquante à deux cent mille individus environ parlant le lithuanien. Sur une longueur de trente à trente-cinq lieues, le lithuanien occupe la frontière prussienne de l'extrême nord-est ; mais il a disparu de toutes les localités importantes, de Memel, de Tilsit, et on ne le rencontre plus que dans les hameaux.

Le groupe des Lithuaniens russes est beaucoup plus compacte ; on les évalue au nombre d'un million trois cent mille, approximativement. Ils n'atteignent ni Grodno au sud, ni Vilna à l'est ; mais ils en approchent de bien près. Au nord, leur limite est celle du lette, dont nous aurons à parler tout à l'heure. Cette frontière septentrionale du lithuanien est à peu près horizontale et s'étend sur une longueur de plus de quatre-vingt-dix lieues. La localité la plus importante du pays où l'on parle lithuanien est la petite ville de Covno.

Schleicher avait divisé le lithuanien en deux dialectes : le bas-lithuanien ou jémaïte, et le haut-lithuanien. Ces deux dialectes ne correspondaient point, d'ailleurs, à la division des Lithuaniens en Russes et en Allemands : en Prusse, aussi bien qu'en Russie, on parlait au nord le jémaïte, au sud le haut-lithuanien.

Pour Schleicher, la différence des deux dialectes consistait principalement en ce fait que là où le jémaïte conservait les groupes *ti*, *di* devant une voyelle, le haut-lithuanien les changeait en « tch , dj » ; le passage de

l'un des dialectes à l'autre serait d'ailleurs tout à fait graduel (1). Cette division en deux groupes a été vivement attaquée. F. Kurschat, tout en reconnaissant qu'en Prusse, aux environs de Memel, on ne se sert point des sons « tch, dj » qu'emploient tous les autres Lithuaniens, ne pense pas qu'il soit possible d'établir sur un assez grand nombre de faits bien déterminés une semblable division. La langue des environs de Memel présenterait bien quelques particularités ; mais, en somme, on ne pourrait en faire un véritable dialecte (2).

Le système des voyelles lithuaniennes est des plus simples, et l'on peut dire qu'après celui du sanskrit et des vieilles langues éraniennes, c'est le système qui est le plus rapproché de l'ancienne langue commune indo-européenne.

En ce qui concerne les consonnes, nous pouvons signaler, entre autres variations, la substitution des simples explosives non aspirées aux anciennes explosives aspirées : là où le sanskrit dit « gh, dh, bh », le lithuanien dit « g, d, b ». Comme les langues slaves et le zend, il connaît notre « j » français et l'emploie souvent à la place d'un « g » ou d'un « gh » de l'indo-européen commun. On le transcrit par un « z » surmonté d'un point. Le lithuanien, enfin, l'emporte sur le sanskrit et sur presque tous les autres idiomes indo-européens, lorsqu'il s'agit de l'ancienne sifflante *s*. Au lieu de lui substituer, comme l'ont fait presque tous les idiomes congénères, une série de nouvelles sifflantes, il la garde toujours telle quelle. Il faut reconnaître que c'est là un grand signe d'antiquité.

La déclinaison du lithuanien est parfaitement conservée. Il n'a point perdu les formes du duel, et les

(1) *Handbuch der litauischen Sprache*, t. I, p. 4. Prague, 1856.
(2) *Wœrterbuch der litauischen Sprache*. Première partie, p. 8 Halle, 1870.

désinences de ses cas rappellent presque toujours avec fidélité les désinences organiques. Dans la conjugaison, enfin, il conserve les formes du présent et le futur ; mais, ayant perdu les quatre autres temps organiques indiquant le passé, il s'est créé un prétérit particulier et un imparfait. Le premier se distingue, en principe, du temps présent, par l'emploi d'une autre désinence, le second est un temps composé, formé de la racine principale, à laquelle on ajoute le prétérit du verbe « faire ».

L'accentuation du lithuanien est des plus difficiles et on ne la connaît pas mieux que l'on ne connaît celle de certaines langues slaves. Quant à son orthographe, elle n'est pas encore fixée ; plusieurs systèmes sont en présence : l'un est plutôt phonétique, l'autre est plutôt étymologique. Tous deux, sans doute, ont leurs avantages particuliers et il serait bien difficile de les concilier.

On possède un monument important de la littérature lithuanienne ; c'est le poème des *Saisons* de Donalitius, en trois mille vers, publié par Rhesa, avec traduction allemande, en 1818 ; par Schleicher, à Pétersbourg, en 1865 ; par Nesselmann, en 1869. Donalitius, né en 1714, mort en 1780, composa d'autres poésies que ses *Saisons;* on possède une partie de ces autres œuvres, et cet ensemble constitue la littérature lithuanienne. L'on a recueilli, en outre, un certain nombre de chants populaires, connus sous le nom de « dainas », un certain nombre de proverbes et de contes en prose. Il y a là un matériel plus que suffisant pour ceux qui veulent étudier ce précieux idiome, dont les jours sont comptés, il est vrai, mais qu'il faudra toujours citer comme un des exemples les plus curieux de conversation linguistique.

II. *Lette.*

On évalue de neuf cent mille à un million et plus le nombre des individus parlant la langue lette. La fron-

tière septentrionale du lithuanien forme sa frontière méridionale ; à l'est, il confine au russe ; au nord, il rencontre une langue ouralo-altaïque, l'ehste. Il occupe le nord de la Courlande, le sud de la Livonie, l'ouest de la province de Vitebsk, et ses centres principaux sont Riga et Mitau.

La grammaire lette est essentiellement la même que celle du lithuanien ; nous ne nous étendrons donc pas sur ce sujet. Il est bon d'ajouter, toutefois, que les formes du lette sont moins bien conservées, en général, que celles du lithuanien. Le lette ne provient certainement pas de cette dernière langue ; mais son caractère général est bien moins antique, bien moins correct. Comme beaucoup de langues qui n'ont point d'autre littérature, le lette possède un certain nombre de chants populaires.

II. *Vieux Prussien.*

Il disparut il y a environ deux cents ans, dans la seconde moitié du dix-septième siècle ; il occupait la côte maritime de la Baltique, de l'embouchure de la Vistule à celle du Niémen. L'allemand a conquis tout l'ancien territoire prussien : ses anciens habitants durent céder peu à peu devant la féodalité et le christianisme qui les envahirent brutalement au treizième siècle et employèrent à cette conquête les dernières violences.

En 1561, le catéchisme allemand fut traduit en prussien ; c'est un des monuments les plus importants que l'on possède pour l'étude de cette langue. Il n'est cependant pas le plus ancien ; Nesselmann a publié, il y a quelques années, un lexique allemand-prussien contenant un peu plus de huit cents mots et qui date du commencement du quinzième siècle.

Moins incorrect que ne l'est souvent le lette actuel, le vieux prussien se rapproche plutôt du lithuanien. Ses

formes sont peut-être moins antiques que celles de ce dernier, mais parfois, cependant, il le surpasse lui aussi. Il dit, par exemple, *nevints* « le neuvième », tandis que le lithuanien, changeant en « d » la nasale organique de ce mot, dit *devintas*.

Le groupe des langues lettiques se rapproche beaucoup, sans doute, du groupe des langues slaves ; on croit ordinairement qu'à une certaine époque langues slaves et langues lettiques étaient réunies en une seule et même forme d'où elles devaient procéder, par la suite, les unes et les autres. Ce qu'il faut penser de cette théorie, nous le dirons un peu plus bas. Quoi qu'il en soit, c'est un fait incontestable que celui de la grande ressemblance de ces deux branches de la famille indo-européenne. Cette ressemblance est si grande, qu'elle a pu tromper bien des personnes et, dans un certain nombre d'écrits ethnographiques, on a classé le lithuanien parmi les langues slaves. C'est là une erreur complète : les langues lettiques et les langues slaves, toutes rapprochées qu'elles sont, n'en demeurent pas moins parfaitement distinctes, comme l'étaient, par exemple, le sanskrit et le perse.

§ 9. Langues indo-européennes non classées.

La plus grande partie des langues indo-européennes, aussi bien des langues mortes que des langues actuellement vivantes, ont été rapprochées de telles ou telles autres langues de la même famille, groupées et classées avec elles. L'on n'est toujours que trop disposé à précipiter les classifications ; la trop grande hâte cependant y est plus nuisible qu'avantageuse, et mieux vaut ne point classer du tout, nous semble-t-il, que classer à la légère, après un premier et superficiel examen. Bopp, lui-même, ne résista pas toujours à ce fâcheux entraînement ; il tenta, à un moment, de classer les langues

du Caucase et les langues maléo-polynésiennes dans le groupe des langues indo-européennes. Cette entreprise n'eut point un heureux succès ; elle montra au moins combien il est difficile, même aux meilleurs esprits, aux esprits les plus critiques, de ne jamais céder à cette tentation.

En traitant, dans notre quatrième chapitre, des langues agglutinantes, nous avons peut-être séparé les uns des autres certains groupes qui dans l'avenir pourront se trouver rapprochés. Nous n'avons pas hésité à présenter comme tout à fait indépendants ces différents groupes.

Il se peut, toutefois, qu'à des signes incontestables on puisse reconnaître qu'une langue appartient, d'une façon générale, à telle ou telle famille d'idiomes ; mais qu'on ne puisse déterminer la place particulière qu'elle occupe dans cette famille, qu'on ne puisse, en un mot, la classer dans aucun groupe ou affirmer qu'elle forme par elle-même, par elle seule, une branche spéciale de cette famille.

Tel est le cas de plusieurs langues indo-européennes éteintes ou encore vivantes, par exemple l'albanais. Nous avons à parler actuellement de quelques-uns de ces idiomes non classés.

I. *Etrusque.*

Peu de langues ont exercé autant que l'a fait l'étrusque la sagacité des grammairiens. Peu de langues aussi ont prêté davantage aux théories les plus opposées et les moins scientifiques. Au quinzième siècle déjà, on faisait descendre l'étrusque de l'hébreu et du chaldéen, et certains auteurs, aujourd'hui encore lui donnent d'une façon générale une origine sémitique, sinon une origine particulièrement hébraïque.

C'est à Lanzi que remonte l'opinion, communément adoptée aujourd'hui, que l'étrusque est une langue italique au même titre que le latin, l'osque, l'ombrien ; le célèbre ouvrage de Lanzi parut en 1789. Cet ouvrage, malheureusement, était dépourvu de méthode, ce qui tenait à la date même de son apparition : lorsqu'il parut, la grammaire comparée des langues indo-européennes n'était pas encore fondée. Lanzi, d'ailleurs, n'avait point les ressources considérables de cette foule d'inscriptions découvertes après lui, et qui constituent maintenant un matériel des plus riches.

Corssen a entrepris de réunir, dans un ouvrage fort important, les résultats acquis à ce jour par les auteurs qui ont traité cette question avec critique et parmi lesquels sa propre place semble marquée (1). La langue étrusque serait, d'après lui, une langue italique, une sœur du latin, de l'osque, de l'ombrien. On aurait déjà reconnu la forme de presque tous les cas, un certain nombre de formes verbales et de formes pronominales. Presque toutes les inscriptions étrusques sont des inscriptions funéraires. On possède un certain nombre d'inscriptions bilingues en latin et en étrusque, trouvées la plupart dans le nord de l'Etrurie, et qui ont été d'un puissant secours, comme il est aisé de le comprendre, pour le déchiffrement de cette langue.

L'alphabet étrusque forme avec l'alphabet ombrien et osque une branche de l'alphabet italique dont nous avons parlé ; il se divise, d'ailleurs, en plusieurs espèces distinctes. Corssen les étudie successivement dans l'ouvrage dont nous avons parlé un peu plus haut. On peut aussi, à ce sujet, consulter les écrits de Conestabile (*Iscrizioni etrusche e etrusco-latine*).

L'origine sémitique de l'étrusque (Stœckel, 1858) doit

(1) *Ueber die Sprache der Etrusker*, Leipzig, 1874, 1875.

être abandonnée. On a cherché (Deecke) à le rattacher aux langues finnoises, mais cela sans succès. Fr. Müller estime que cette langue est — comme celle des Basques — la langue d'un peuple ancien qui a disparu et dont l'origine demeurera toujours inconnue. Fligier (*Zur præhistorischen Ethnol. Italiens*, 1877), est d'avis qu'à ce sujet nous ne savons rien de positif, et il partage cette opinion de Denis d'Halicarnasse, que les Etrusques sont isolés de tous les autres peuples par leur langue et leurs mœurs.

Il convient, en somme, d'être fort réservé et d'attendre le résultat de nouveaux travaux. Notre opinion est que Corssen n'a point démontré que l'étrusque fût une langue italique ; mais on peut supposer — sans toutefois rien affirmer encore — que c'est un idiome indo-européen. Il appartient à l'avenir de résoudre le problème.

II. *Dace.*

La Dacie ancienne, limitée au sud par le Danube, au nord-est par le Dniester, au nord-ouest par la Tisza, comprenait les régions qui forment aujourd'hui le cercle hongrois d'au-delà de la Tisza, la Transylvanie, la Bucovine, le banat de la Temes, la Valachie, la Moldavie et la Bessarabie.

Il ne reste de la langue dace que bien peu de débris, quelques noms de plantes cités par le médecin Dioscoride et un certain nombre de noms géographiques. Ces noms ont incontestablement une apparence indo-européenne; *propedula* « quintefeuille » rappelle la forme celtique *pempedula*. Mais le dace était-il une langue celtique, une langue germanique, une langue slave ? appartenait-il à quelque autre branche de la famille indo-européenne ? formait-il par lui-même une branche indépendante et

distincte de toutes les autres ? C'est ce que l'on ne peut décider en l'état de la question.

Un écrivain roumain, Hasdeu, explique sans aucune hésitation tous les noms géographiques daces qui nous ont été conservés par Ptolémée, Strabon, la Table de Peutinger; bien plus, il a cru retrouver l'ancien alphabet dace dans un alphabet qui s'était conservé, jusqu'au siècle dernier, chez les Széklers de la Transylvanie. Pour lui le dace aurait appartenu à une famille thraco-illyrienne à laquelle se rattacheraient, entre autres idiomes, le phrygien et l'albanais (1). Cette thèse aurait demandé plus de développements que ne lui en a accordé son auteur. Il eût été bon de discuter l'opinion de ceux qui ont rattaché le dace aux langues germaniques (Grimm), aux langues slaves (Müllenhoff), ou encore aux langues celtiques. (Consulter Dieffenbach, *Vœlkerkunde Osteuropas*, t. Ier, p. 124.)

III. *Langues indo-européennes de l'Asie Mineure.*

Il paraît avéré, aujourd'hui, qu'une grande partie des langues de l'Asie Mineure appartenaient à la famille des langues indo-européennes (2). C'est incontestablement le cas du phrygien et du lycien.

On possède un assez grand nombre d'inscriptions lyciennes, dont quelques-unes sont bilingues, grec et lycien. Cette dernière circonstance facilitera grandement, sans aucun doute, les progrès du déchiffrement de cette langue; on peut dire, d'ailleurs, que son alphabet est fixé à présent d'une façon à peu près certaine.

Du phrygien on possède également quelques inscriptions, trouvées en Phrygie même, et une série de mots

(1) *Istoria critiqua a Romaniloru.* Deuxième édition, t. I, p. 292. Bucharest, 1874.
(2) Renan, *Histoire des langues sémitiques*, liv. I, chap. II, § 2.

cités par les auteurs anciens. Le nombre de ces derniers mots est assez important, et comme leur sens est bien fixé par les auteurs mêmes qui les rapportent, ils peuvent servir de point de départ, à toute l'étude du phrygien. Leur inscription, sans doute, peut être plus ou moins exacte, mais il ne faut pas penser qu'elle soit trop défectueuse.

Que l'on rapproche ces langues indo-européennes du grec ou des langues éraniennes, notamment de l'arménien, la transcription des mots de ces différents idiomes en langue grecque doit être relativement fidèle. Les anciennes langues éraniennes, en effet, ne sont pas fort éloignées des dialectes grecs, et il est permis de penser que les idiomes indo-européens de l'Asie Mineure relient davantage ces deux familles.

Ils n'appartiendraient donc ni au groupe des langues éraniennes, comme l'ont pensé beaucoup d'auteurs, ni au groupe des dialectes grecs, mais ils formeraient une branche spéciale aussi rapprochée du grec que de l'arménien et de l'ancien perse.

Ce n'est encore là qu'une simple hypothèse que l'avenir pourra tout aussi bien renverser ou confirmer. Peut-être découvrira-t-on, d'ailleurs, que si certains idiomes de l'Asie Mineure sont intimement alliés, comme, par exemplé, le coréen et le lycien, il en est d'autres qui n'ont entre eux que des rapports assez éloignés ; peut-être même faudra-t-il les classer en deux groupes, dont l'un se rattacherait aux langues éraniennes, l'autre au grec. Mais la question est dans la première période d'étude, et ces différentes langues ne peuvent être rangées que parmi celles dont la classification n'est point encore possible.

IV. *Langues indo-européennes dites « scythiques ».*

Au paragraphe dix-neuvième de notre quatrième chapitre nous avons dit que les expressions de « Scythes » et de « scythique » n'étaient que des noms géographiques et qu'elles s'appliquaient à un grand nombre de populations ; différentes de race et de langue. Nous avons dit également que certaines populations appelées « scythiques » par les auteurs anciens, parlaient un idiome indo-européen (1) ; le lecteur voudra bien se rapporter à ce passage et nous ne mentionnons le fait, ici, que pour mémoire. Consulter, d'ailleurs, Fligier, *Zur Skythenfrage*, Vienne, 1878.

V. *Albanais.*

La question de l'origine de l'albanais et son classement dans la famille indo-européenne ont tourmenté bien des linguistes ; le problème n'est pas encore résolu.

L'albanais occupe la région de l'empire turc donnant sur la mer Adriatique, la passe d'Otrante et la mer Ionienne. Au nord, le territoire albanais confine aux Serbes monténégrins et à ceux qui font partie intégrante de l'empire ; à l'est, il confine dans sa partie supérieure aux Bulgares et dans sa partie inférieure aux Grecs de l'empire turc ; au sud, aux Grecs également. La plus grande longueur de ce territoire est d'environ quatre-vingt-quatre lieues, sa largeur moyenne de trente lieues environ. Au nord-est de Scutari, il comprend des enclaves serbes assez importantes, au centre, et surtout au sud, à l'est de Janina, des enclaves arméniennes non moins considérables.

(1) Girard de Rialle, Bulletins de la Société d'anthropologie de Paris, 1869, p. 46.

Le nombre des Albanais est d'environ un million et demi d'individus ; beaucoup moins nombreux que les Slaves de Turquie, ils l'emportent d'autre part sur les Turcs eux-mêmes et sur les Grecs soumis encore à la domination ottomane. Le nom véritable de l'albanais est celui de « skipetar ».

Quelques auteurs ont voulu rapprocher l'albanais des langues slaves ; cette tentative a toujours échoué et il n'est point vraisemblable qu'elle réussisse jamais. Une opinion plus répandue a considéré l'albanais comme un parent assez intime de la langue grecque (Hahn Camarda), mais on peut dire que cette assertion n'a jamais été vérifiée scientifiquement. D'après une troisième opinion (Blau), l'albanais doit être rattaché aux idiomes éraniens. Ces deux dernières opinions se concilieraient d'ailleurs assez bien si l'on admettait que les idiomes éraniens et le grec sont fort rapprochés l'un de l'autre (Picot). D'autres auteurs ont cherché à établir un rapprochement plus ou moins intime entre l'albanais et les langues italiques. Sont-ils plus près de la vérité que les partisans d'une origine hellénique ? C'est ce que nous ne voulons pas décider. La question, à nos yeux, demeure encore tout entière à résoudre. On sait que l'adjectif albanais possède, comme celui des Slaves, une sorte d'appendice d'origine pronominale, que le nom se suffixe un article, comme font le roumain ou le bulgare, mais tout le reste est fort obscur, surtout la conjugaison.

Ce qui rend l'étude de l'albanais particulièrement difficile, c'est que le lexique de cette langue est en grande partie composé d'éléments étrangers, latins, grecs, slaves, turcs et autres. On parviendra peut-être à dégager ces éléments d'emprunts. Déjà F. Miklosich a dressé un tableau des mots tirés du latin et des langues slaves ;

ces derniers sont en grande partie des mots que le roumain a empruntés lui aussi.

Jusqu'à preuve nouvelle, pensons-nous donc, l'albanais ne peut que passer purement et simplement pour une langue indo-européenne ; ce fait est bien acquis, mais on ne saurait guère aller plus loin et rattacher d'ores et déjà l'idiome en question à telle ou telle branche particulière du groupe indo-européen.

§ 10. Du mode de subdivision de la langue indo-européenne et de la région où elle fut parlée.

I

A peine avait-on constaté la parenté des différentes langues indo-européennes, à peine avait-on reconnu qu'elles descendaient toutes d'un ancien idiome dont l'histoire avait perdu les traces, que l'on songea à les classer entre elles. Il s'agissait de les grouper selon leur degré d'affinité, de les réunir en familles et de rapprocher à leur tour les unes des autres les familles qui paraissent offrir des traces d'une parenté plus intime. En d'autres termes, il s'agissait de diviser la souche indo-européenne en branches, ces branches en rameaux, et ainsi de suite.

Le premier rapprochement que l'on établit fut celui du grec et du latin ; on y était inévitablement poussé par les traditions de la philologie classique.

On supposa donc qu'une seule et même langue, détachée des autres idiomes indo-européens, avait donné naissance à deux langues sœurs, à deux langues jumelles, le grec et le latin. Cette branche gréco-latine, à laquelle il parut opportun de donner un nom, reçut celui de « pélasgique ». Jamais appellation ne fut moins justi-

fiée. Loin de savoir, en effet, ce que c'étaient que les Pélasges, on peut à peine assurer qu'un peuple quelconque ait en aucun temps répondu à ce nom, et les quelques passages des histoires d'Hérodote où il se trouve relaté suffisent à empêcher tout auteur sérieux de lui attribuer une acception déterminée.

Les travaux d'Eugène Burnouf et de Lassen sur l'ancien perse et le zend permirent de rapprocher intimement les langues éraniennes du sanskrit. On supposa donc qu'il avait existé une langue commune indo-éranienne dont le sanskrit, d'une part, et les langues éraniennes, d'autre part, seraient sortis à un moment donné.

La grande ressemblance du lithuanien et des langues slaves fit accepter également une langue commune letto-slave ; cette langue letto-slave aurait, à son tour, une origine commune avec le type des langues germaniques, et ainsi de suite.

Plusieurs systèmes assez tranchés se trouvent ici en présence. Certains auteurs, par exemple, ont adopté le tableau que voici :

- Indo-européen.
  - Indo-éranien.
    - Sanskrit.
    - Eranien.
  - Européen.
    - Gréco-italique.
      - Grec.
      - Italique.
    - Celtique.
    - Germano-letto-slave.
      - Germanique.
      - Letto-slave
        - Lettique.
        - Slave.

Schleicher envisageait cette répartition d'une façon différente et dressait cet autre tableau :

- Indo-européen.
  - Letto-slavo-germanique.
    - Germanique.
    - Letto-slave.
      - Lettique.
      - Slave.
  - Aryo-gréco-italo-celtique.
    - Gréco-italo-celtique.
      - Italo-celtique.
        - Celtique.
        - Italique.
      - Grec.
    - Arique.
      - Eranien.
      - Hindou.

Dans ce tableau, il n'y a donc plus de langue spécialement européenne, et une partie des langues de l'Europe seraient plus rapprochées du sanskrit et des idiomes éraniens que des autres langues européennes. Cette théorie, malgré l'autorité de son auteur, ne paraît pas avoir gagné beaucoup de partisans. Généralement on a préféré s'en tenir à la division en indo-éranien et en européen (1). Certains auteurs, admettant d'ailleurs cette double division, comprenaient de différentes façons les sous-divisions ; les uns, par exemple, rapprochaient davantage les langues celtiques des langues germaniques, d'autres les rapprochaient plus volontiers du latin.

D'ailleurs, la théorie de la ramification de la souche commune indo-européenne n'est pas acceptée universellement. Elle a été attaquée simultanément en France et en Allemagne dans deux écrits tout à fait indépendants l'un de l'autre et publiés isolément à la même époque. L'un de ces écrits est de l'auteur de ces lignes (2), l'autre de Johannes Schmidt (3). Cet auteur admet encore une

(1) Havet. *L'Unité linguistique européenne*, Mémoires de la Société de linguistique, t. II, p. 261.

(2) *Notice sur les subdivisions de la langue commune indo-européenne*. Comptes rendus de la première session de l'Association française pour l'avancement des sciences, p. 736. Bordeaux, 1872.

(3) *Die Verwandtschaftsverhältnisse der indo-germanischen Sprachen*. Weimar, 1872.

unité linguistique indo-éranienne et une unité letto-slave, mais il se refuse à aller plus loin. Il cherche à démontrer que, si du côté de l'occident les langues slaves et lettiques sont indissolublement liées aux langues germaniques, elles se trouvent tout aussi liées, du côté de l'orient, aux idiomes éraniens et hindous : non seulement, donc, il n'a point existé d'idiome commun germano-letto slave, mais il n'a point existé non plus d'idiome spécialement européen, nettement distinct du sanskrit et des langues éraniennes. Le grec, d'autre part, serait tout aussi inséparable d'avec les deux familles asiatiques que d'avec la branche italique, et les langues celtiques ne pourraient pas être groupées à plus juste titre avec les langues italiques qu'avec les langues germaniques. Cette question n'est pas de celles que l'on puisse trancher après une étude de quelques instants, car elle est fort complexe.

En ce qui nous concerne, nous pensons qu'il n'a point existé de groupes secondaires entre la langue commune indo-européenne et les groupes éranien, hellénique, germanique et autres.

Sans doute, certains idiomes indo-européens sont plus rapprochés, en somme, de quelques-uns de leurs congénères que de quelques autres d'entre eux ; le latin, par exemple, est plus intimement allié aux langues celtiques qu'aux langues éraniennes. Mais s'ensuit-il qu'il faille conclure à une langue commune italo-celtique ? Assurément non.

Nous ne connaîtrons jamais, selon toute vraisemblance, les motifs qui déterminèrent les populations dont la langue était l'indo-européen commun à entreprendre leurs grandes migrations ; mais nous pouvons penser, sans crainte d'erreur, qu'avant leurs migrations, ces populations occupaient un territoire assez vaste. En ces larges limites la langue commune indo-européenne ne devait-elle point se modifier, s'altérer, se corrompre de

façon différente dans les différentes tribus établies sur ce territoire ? Nous pensons qu'il n'en pouvait être autrement. Ces modifications, ces altérations ne furent évidemment pas les mêmes en tous lieux ; ici, par exemple, elles purent s'attaquer de préférence aux sifflantes, là aux explosives, ailleurs aux formes elles-mêmes des mots. On peut admettre en outre que, selon toute vraisemblance, les modifications qu'acceptait telle ou telle tribu devaient, à peu de chose près, être de la même nature que les modifications acceptées par la tribu voisine ; plus les groupes se trouvaient distants, plus ils devaient montrer de différences. En d'autres termes, il devait y avoir plus de diversité entre le groupe de l'extrême est et celui de l'extrême ouest, qu'entre ce dernier et un groupe central. Cette espèce de série, cette sorte de continuité est toute naturelle et, de nos jours, nous la retrouvons dans les patois.

Nous n'avons pas à nous enquérir des causes qui déterminèrent la tendance générale propre à tel ou tel ensemble de tribus voisines. Ces causes nous resteront peut-être à jamais inconnues ; mais ce que nous pouvons parfaitement admettre, c'est que ces unités secondaires, ces branches intermédiaires dont nous parlions tout à l'heure, par exemple la prétendue langue italo-celtique, n'ont jamais eu d'existence réelle. C'est un besoin immodéré de classification qui les a mises au jour. En fait, elles n'ont point vécu. On les a multipliées, mais on pouvait les multiplier bien plus encore ; il serait facile de restituer un idiome commun helléno-slave, érano-celtique, italo-germanique. Une fois dans le domaine de la fantaisie, il n'y aucun motif de s'arrêter.

II

Avant d'en terminer avec les langues indo-européennes, il nous reste à dire quelques mots d'une question assez débattue et que nous ne pouvons passer sous silence, celle du pays où fut parlée la langue indo-européenne commune.

Distinguons tout d'abord la question de race de la question de la langue.

Lorsqu'il s'agit de la formation même du langage articulé, le facteur de la race est non seulement capital, il est unique. Acquisition de la faculté du langage articulé, formation des premiers systèmes de langues et formation des premières races humaines tombent à un seul et même moment. Cela a été le sujet de notre chapitre second, sur lequel nous n'avons point à revenir en cet instant. Nous insistons simplement sur ce fait que, si les races d'Europe sont réellement des races européennes, telles, au moins, que nous les voyons aujourd'hui, il ne s'ensuit en aucune façon que les langues indo-européennes de nos régions y aient également pris naissance. Cette distinction est capitale, et on la néglige trop souvent. Nous pouvons dire plus encore ; s'il est juste de parler de langues indo-européennes, il est absolument vicieux de parler d'une « race » indo-européenne. Une telle race n'existe point, et ceux-là seuls peuvent en disserter, la décrire même et tracer ses frontières, qui n'ont aucune notion d'anthropologie.

Allons plus loin. S'il est certain qu'une langue indo-européenne commune a été parlée jadis en une région quelconque, il n'est nullement certain que les individus parlant cette langue aient appartenu à une seule et même race. L'indo-européen commun a été formé, sans doute, en un seul et unique centre ; mais la période de formation une fois passée, rien ne dit qu'il ne se soit pas

étendu sur différentes populations très étrangères, comme nous avons vu le latin rustique s'étendre sur les populations voisines du Guadalaviar, de la Garonne, de la Somme, de l'Adige et du bas Danube. Bien des hypothèses sont permises à ce sujet. En définitive, il n'y a ici qu'un fait, un seul fait bien avéré auquel nous puissions nous en tenir : le fait de l'existence de cette langue commune indo-européenne, abstraction faite de toute question de race.

Cela bien acquis, nous pouvons aborder notre sujet sans crainte de malentendus.

Il y a une trentaine d'années, l'on s'accordait assez généralement à donner pour patrie à la langue commune indo-européenne « le vaste plateau de l'Iran », comme dit Pictet, « cet immense quadrilatère qui s'étend de l'Indus au Tigre et à l'Euphrate, de l'Oxus et du Iaxartes au golfe Persique » (1). Cette région correspond à la Perse actuelle et aux pays qui lui sont limitrophes à l'est et à l'ouest. On trouva, avec juste raison, que l'indication d'une aire aussi considérable était trop vague, et l'on chercha à la restreindre. Les traditions de l'Avesta aidant, on émit l'opinion que la Baktriane devait être regardée comme la patrie ancienne des prétendus « Aryas », c'est-à-dire des individus qui parlaient la langue commune indo-européenne. Mais n'était-ce pas donner à la tradition éranienne un sens beaucoup plus large que celui qu'elle avait en réalité ? A la rigueur, l'Avesta pouvait encore se souvenir d'une patrie plus ancienne des Eraniens ; mais celle-ci avait-elle été en même temps la patrie commune de toutes les familles indo-européennes ? On n'était pas en droit de l'affirmer ; pareille conclusion dépassait beaucoup les prémisses. C'est ce que l'on a compris facilement.

(1) *Les Origines indo-européennes ou les Aryas primitifs. Essai de paléontologie linguistique*, t. I, p. 35. Paris, 1859.

Les moyens les plus sûrs d'arriver à la solution cherchée devaient être demandés à la linguistique.

Les indications que peut donner le lexique comparé des langues indo-européennes sur les termes géographiques et topographiques, sur les noms des cours d'eau, des montagnes, sur ceux des métaux, des plantes, des animaux, ne donnent que des renseignements très vagues. On peut les attribuer à une foule de régions ; ils s'appliquent aussi bien à l'Assyrie qu'à la Baktriane.

En s'en rapportant à la physionomie générale des idiomes indo-européens, on conclut du fait de la conservation même du sanskrit, à cette opinion qu'il devait être moins éloigné que ses congénères de l'ancien centre linguistique. Il sembla raisonnable d'admettre que ceux de ces idiomes qui, dans leur ensemble, se rapprochent de la façon la plus fidèle du type indo-européen commun, sont également ceux qui se sont le moins éloignés des régions où ce type commun était parlé.

De là cette première conclusion : entre toutes les langues indo-européennes, le sanskrit et les langues éraniennes sont celles qui se sont le moins éloignées de la région où était parlé l'indo-européen, tandis que les langues celtiques s'en sont éloignées plus que ne l'ont fait toutes les autres.

Au second degré de conservation, l'on dut placer les dialectes grecs au sud-est de l'Europe, les langues lettiques et slaves au nord-est ; au troisième degré, les langues germaniques au nord, les langues italiques au sud. Ces deux dernières branches rejoignent l'une et l'autre les langues celtiques placées au quatrième et dernier degré.

Pictet conclut. Traçant une ellipse assez allongée, il a regardé l'un des foyers de cette ellipse, celui de droite, comme le point où aurait été parlé l'indo-européen commun. A peu de distance de ce foyer, vers la droite, il

place au bas le sanskrit, plus haut les langues éraniennes. Suivant ensuite de droite à gauche les deux branches de l'ellipse, il place au centre, dans le haut, les langues letto-slaves ; au centre, dans le bas, les idiomes grecs ; ces deux groupes sont encore assez rapprochés du foyer de droite, mais moins que ne le sont les langues éraniennes et le sanskrit. Poussant encore vers la gauche, Pictet place les langues germaniques en haut, et les langues italiques en bas, dans la même position vis-à-vis du foyer de gauche qu'occupent les langues éraniennes et le sanskrit vis-à-vis du foyer de droite. Plus à gauche encore, tout à l'extrémité de la ligne transverse horizontale de l'ellipse, se trouvent les langues celtiques, entre les langues germaniques et italiques : elles sont ainsi les langues les plus éloignées du foyer de droite, c'est-à-dire du prétendu point de départ.

Il est aisé de construire cette figure. Elle est sans doute très ingénieuse, et au premier moment on est fort tenté de l'adopter ; elle concorde assez bien, d'ailleurs, avec l'hypothèse qui regarde la Baktriane comme la région où fut parlée la langue indo-européenne.

Mais, en réalité, on peut l'interpréter de deux façons et lui donner deux sens bien tranchés. Le premier sens est celui qu'en tire Pictet ; voici le second : il se peut que le centre commun recherché ne se trouve pas au foyer de droite de l'ellipse, mais qu'il soit situé plus sur la droite, en dehors même de l'ellipse, c'est-à-dire vers la frontière chinoise. Avec cet autre centre, le sanskrit et les langues éraniennes resteraient toujours au premier degré, le grec et le letto-slave au second, les langues germaniques et italiques au troisième, les langues celtiques au quatrième et dernier.

La théorie qui plaçait dans l'Asie centrale, dans la haute Asie, la patrie de l'indo-européen commun a été vivement attaquée.

L'Anglais Latham fut le premier, semble-t-il, qui opina pour une origine européenne. Quelques auteurs l'ont suivi. Il en est parmi ceux-ci qui se sont efforcés de donner quelque apparence scientifique à leur assertion, il en est d'autres qui ont tranché net cette question spéciale avec autant d'audace que d'incompétence.

Certaines personnes, par exemple, voyant les mots celtiques plus courts que les mots sanskrits, en ont inféré qu'ils étaient plus simples, partant plus primitifs, et s'éloignaient moins du type commun. C'est de la linguistique au millimètre. Avec ce procédé, l'anglo-saxon proviendrait de l'anglais, le latin du français, le vieux perse du persan.

D'autres auteurs, arguant de ce fait que le type blond aux yeux bleus se présente plus particulièrement dans des pays de langue allemande, en concluent, on ne sait trop pourquoi, que l'indo-européen commun a été parlé en Germanie. Ils confondent ici la langue et la race, ou, pour mieux dire encore, la langue et les races ; c'est une méprise sur laquelle nous ne pouvons même pas nous arrêter.

Peu nous importe que les populations qui parlèrent l'indo-européen commun aient été blondes ou brunes, ou qu'il y en ait eu parmi elles et de blondes et de brunes ; ce n'est pas en question : la langue seule nous occupe et non point la race. Nous n'appelons même pas à notre aide le secours de l'archéologie, qui, pourtant, enseigne d'une façon claire et nette qu'à l'époque où l'Orient connaissait déjà une certaine civilisation, l'Occident en était encore à l'état sauvage, ou à peu près.

Benfey, se fondant sur le manque de noms primitifs communs pour les grands carnassiers asiatiques, opina que l'indo-européen devait avoir été parlé en Europe (*Geschichte der Sprachwissensch.*, p. 600). Peschel, Brunnhofer et d'autres auteurs opinent pour l'Arménie.

Fr. Müller se prononce pour la demeure européenne, sans la préciser, et est d'avis que les langues aryennes de l'Asie ont gagné par voie de migration les régions où elles sont aujourd'hui parlées. Les grammairiens qui tiennent le grec pour le mieux conservé des idiomes de la famille supposent que la patrie primitive dont il s'agit doit être cherchée vers le sud-est de l'empire. Penka (*Die Herkunft der Aryer*, 1886) fait de la Scandinavie le premier centre de rayonnement des soi-disant Aryens et place leur première demeure dans l'Europe du sud-ouest. Cette opinion ne concorde pas avec les derniers résultats de l'anthropologie. Il semble, au contraire, que les idiomes aryens aient été importés, vers la fin de l'industrie de la pierre taillée, par la population brachycéphale qui, venant de l'est, pénétra jusque vers l'Atlantique. Cette population apporta sa langue, ses mœurs, aux anciens habitants dolichocéphales (Cf. Hovelacque et G. Hervé, *Précis d'anthropologie*, p. 367, 583). L'erreur de Penka provient de ce qu'il confond les dolichocéphales « paléolithiques » avec les dolichocéphales « néolithiques »,, ancêtres des blonds de l'Allemagne du Nord.

Si toutefois l'on remonte le courant de l'invasion de la population à tête courte, que nous supposons avoir apporté dans l'Europe occidentale les langues indo-européennes, on s'arrête, après avoir franchi l'Europe centrale, aux pays du bas Danube. Faut-il remonter plus loin vers l'est ? C'est ce qu'il est difficile de faire dans l'état actuel des connaissances. Cette question vidée, il restera à rechercher dans quelle région les blonds du nord de l'Europe ont reçu eux-mêmes les idiomes indo-européens.

En somme, l'obscurité est loin encore d'être dissipée.

CHAPITRE VI.

PLURALITÉ ORIGINELLE DES LANGUES ET TRANSFORMATION DES SYSTÈMES LINGUISTIQUES.

Arrivés au terme de ce long examen, nous avons à jeter en arrière un coup d'œil d'ensemble et à récapituler, dans un chapitre final, les points les plus importants de notre étude.

Nous avons tout d'abord à revenir sur la question de la doctrine et de la méthode. Ce sujet est le premier qu'il nous ait fallu aborder ; c'est celui qui doit attirer en dernier lieu notre attention. La doctrine, la méthode dominent les sciences contemporaines ou, pour mieux dire, elles font corps avec elles, et l'on ne saurait trop insister sur ce fait capital ; cette alliance indissoluble, cette sorte d'identification est la caractéristique même de ce nouvel état de choses.

§ 1. Comment se reconnaît la parenté des langues.

Il arrive souvent aux personnes dont les connaissances linguistiques ne sont que très imparfaites, très superficielles, de regarder sans hésitation comme proches parentes les familles de langues qu'un auteur vraiment compétent n'osera point rapprocher les unes des autres et que parfois même il déclarera irréductibles. L'étymologie n'est jamais plus dangereuse que sur ce terrain. Pour tout dire, elle n'y connaît pas d'obstacles.

Les préoccupations bibliques ont contribué plus que toute autre cause à développer la manie funeste de l'étymologie. Il s'agissait, il importait de rattacher aux langues sémitiques tous les idiomes de l'univers, ou bien par voie de descendance directe, ou bien par voie de parenté collatérale. On renonçait, au besoin, à faire de l'hébreu la mère de toutes les langues, mais il fallait au moins les rattacher toutes, y compris l'hébreu, à une souche commune à une seule et même langue mère.

C'est un fait qui ne se laisse plus discuter à l'heure actuelle, et, sous quelques réserves que ce soit, parler encore de cette langue commune soi-disant primitive, c'est faire preuve d'une complète ignorance de la méthode linguistique.

Avant tout, dans la comparaison des langues, il faut négliger la ressemblance pure et simple des mots. Deux mots dont le sens est presque le même dans deux langues différentes, dont le sens, si l'on veut, est absolument le même, peuvent n'avoir rien de commun. La concordance lexique sans la concordance grammaticale est nulle et non avenue. L'étymologiste s'en empare, s'en contente et ne veut pas voir plus loin ; le linguiste ne s'y arrête même pas.

Aux yeux de ce dernier, la dissection de deux mots plus ou moins semblables peut seule prouver leur parenté ; à aucun titre il ne s'accorde le droit de comparer deux mots tout faits. Les éléments formatifs de deux mots sont-ils bien les mêmes, leur racine est-elle également la même, dans ce cas il est légitime de les regarder comme deux mots correspondants, de leur donner une origine commune ; si ces conditions ne se trouvent pas réunies, les deux mots en question ne peuvent être identifiés, quelle que soit leur homophonie.

Prenons pour exemple le mot *un* dans différentes langues et voyons comment on a pu rapprocher sans cri-

tique aucune les mots qui servent à l'exprimer dans les langues en question.

Tout d'abord les celtomanes n'ont pas manqué de trouver au mot français *un* une origine celtique, grâce au gallois *un*, au cornique *un*, à l'armoricain *eun*. A leurs yeux, ces formes diverses se rapprochent plus du mot français dont il s'agit que ne s'en rapproche le latin *unus : un* celtique, *un* français c'est, disent-ils, un seul et même mot ; c'est du celtique et non du latin, que provient donc le français. Rien n'est plus inexact. Le celtomane, en effet, néglige ici deux facteurs importants. L'un de ces facteurs, c'est la forme antique du français *un*. Au onzième siècle, dans la langue française à deux cas, le nominatif du mot en question était *uns*. Est-ce le celtique *un*, *eun* qui expliquera la sifflante qui se trouve à la fin de ce nominatif *uns ?* Assurément non. C'est le latin *unus* qui pourra seul en rendre compte. Un autre facteur, avons-nous dit, a été également négligé dans le rapprochement du mot celtique et du mot français. Avant de parler d'un celtique *un*, il s'agissait de comparer le gallois et le cornique *un* au gaélique *óin* et il fallait ramener ces deux formes à une forme commune. Mais le celtomane n'a que faire des procédés méthodiques de la linguistique. Le celtomane n'est qu'étymologiste, et s'il n'était étymologiste il ne serait pas celtomane.

N'a-t-on pas rapproché aussi l'anglais *one* « un » et le malayâla *onn' ?* Il suffisait cependant, pour éviter ce rapprochement imprévu, de comparer d'une part la forme du malayâla aux autres formes dravidiennes, et de se rappeler le thème gothique *aina-*, au nominatif *ains*.

Nous devons beaucoup sans doute aux missionnaires qui songent à rapporter des pays peu connus qu'ils ont visités des séries de mots et de phrases, des essais de

grammaire, ces essais fussent-ils (comme c'est presque toujours le cas), absolument dépourvus de méthode ; mais que dire de la manie étymologique qui les tourmente à peu près tous sans exception ? Ils n'hésiteront jamais, par exemple, à comparer des mots polynésiens, des mots cafirs, des mots américains à des mots latins ou français. Que le français et le latin aient une histoire, un long passé, cela leur est complètement indifférent ; que l'américain soit agglutinant, que le latin possède une flexion véritable, cela leur importe peu. Ils prennent les mots tout faits, au hasard, sous leur forme actuelle et décident sans plus tarder de leur identité.

L. Adam a très justement dit, au premier Congrès des américanistes, où l'on rapprochait sans critique aucune le basque, le bas-breton, l'algonquin et bien d'autres langues encore : « Dans l'intérieur d'une même famille les rapprochements de mots sont légitimes et concluants, à la condition d'être opérés en conformité avec les règles de la phonétique et de la dérivation, sans le respect desquelles l'étymologie n'est qu'un art puéril, indigne d'occuper l'attention des vrais savants. Quand, après de fortes études préparatoires, un linguiste sachant son métier aborde la texicologie d'une famille de langues, qu'il se renferme dans ce domaine et qu'il opère scientifiquement, c'est-à-dire d'après des règles certaines, les rapprochements qu'il fait ont toujours chance d'être fondés. Que si, au contraire, il entreprend de passer d'une famille à une autre, ni la science acquise, ni les règles ne lui serviront à rien, et il aboutira fatalement à des résultats sans consistance. » *Op. cit.*, t. II, p. 40.

C'est en termes non moins explicites que J. Vinson répondit à ces mêmes rapprochements du basque et des langues américaines : « Pour déterminer la nature et la place naturelle d'un idiome nouveau, le linguiste doit tenir compte des particularités qu'il présente dans cha-

cune des divisions de la grammaire. Il faut, pour qu'une langue soit définitivement classée, connaître les sons qu'elle emploie et leurs variations, les éléments formatifs dont elle se sert et leur mode de groupement, les racines qui constituent son corps matériel, enfin les règles principales de sa syntaxe. Il n'est pas moins nécessaire de ne comparer que des idiomes pris à un même degré de formation, en les ramenant par exemple au point culminant de leur developpement formel. Enfin, pour conclure à une communauté d'origine de deux idiomes, il serait indispensable que leurs principaux éléments grammaticaux soient non seulement analogues par leur fonction, mais encore qu'ils se ressemblent phonétiquement d'une manière suffisante pour rendre admissible l'hypothèse de leur identité primitive.

« La parenté de deux ou plusieurs langues ne saurait en effet résulter uniquement d'une même physionomie extérieure. Si les racines significatives, qui sont, après tout, le fond propre, la haute originalité du langage, se trouvent totalement différentes de l'une à l'autre, il sera sage de ne point affirmer que ces langues proviennent d'une source commune... Que prouvent des listes de mots réunis sans ordre par un voyageur, un amateur de circonstance, qui n'a d'autre mérite, d'autre expérience, d'autre science même que sa bonne volonté ? Pour que de pareils rapprochements soient probants, il faut qu'ils viennent seulement après que l'on a démontré l'identité générale des grammaires, après qu'on a distingué les éléments formatifs, après qu'on a ramené les mots significatifs et les mots de relation à leur plus simple et plus primitif aspect sonore. » *Op. cit.*, t. II, p. 52 ss.

On ne saurait trop le répéter, serait-ce des centaines et des centaines de mots tout faits appartenant à deux langues quelconques qu'on eût à comparer, cette comparaison ne ferait point avancer d'un pas la question

de la parenté de ces deux langues. Ce qu'il s'agit de démontrer, c'est tout autre chose que l'existence de ces rapports forfuits : c'est l'identité des racines réduites à leur forme la plus simple ; c'est l'identité des éléments formatifs du mot ; c'est l'identité de fonctionnement de ces éléments; en un mot, c'est l'identité grammaticale.

Il n'y a point à tenir compte des études soi-disant comparatives qui ne seraient pas basées sur ces principes sévères ; elles ne sont plus de notre temps.

Ce n'est pas, d'ailleurs, qu'il soit toujours facile de ne point se laisser entraîner. Bopp lui-même a tenté par deux fois des rapprochements bien malheureux entre des groupes de langues tout à fait distincts. Les langues du Caucase lui parurent un jour se rattacher aux langues indo-européennes, et il chercha à le démontrer. Ce fut sans succès. Dans cette malheureuse tentative, Bopp n'était pas demeuré fidèle à la méthode qui lui avait si bien réussi dans la comparaison des langues indo-européennes.

Il ne fut pas plus heureux lorsqu'il s'efforça d'apparenter aux mêmes langues indo-européennes, notamment à la branche hindoue, les langues maléo-polynésiennes. On a sans peine démontré scientifiquement combien ces deux groupes d'idiomes sont à tous égards différents l'un de l'autre.

Tout d'abord leur phonétique est parfaitement distincte ; le malai, par exemple, ne connaît pas les aspirées dont le sanskrit et les langues néo-hindoues sont si bien pourvues. A l'encontre du sanskrit, le malai peut dériver la racine au moyen de préfixes, c'est-à-dire d'éléments placés avant, et non après cette racine ou le thème du mot : dans les langues indo-européennes, la dérivation a toujours lieu par suffixes, c'est-à-dire par les éléments placés après la racine, après le thème, et non avant. Le malai ne connaît point, comme les langues

indo-européennes, des cas véritables, des suffixes personnels véritables ; et nous pouvons ajouter que ces différences capitales ne sont pas les seules (1).

Bopp s'était fié malheureusement à des rapprochements lexiques (qu'expliquaient d'ailleurs fort bien de nombreux emprunts faits à l'Inde par les Malais), et cette sorte de rapprochements, nous ne saurions trop le répéter, n'a de valeur qu'autant qu'ont été résolues la question de la phonétique et celle de la formation même des mots.

§ 2. Pluralité originelle des systèmes linguistiques.

On a cherché souvent à rapprocher l'un de l'autre le système des langues indo-européennes et celui des langues sémitiques : on a voulu leur trouver une origine commune, les ramener à des formes communes. On a toujours échoué dans cette entreprise. Le sanskrit est aussi distinct de l'arabe et de l'hébreu qu'il l'est du tagala et du javanais. Non seulement il n'y a point d'identité grammaticale entre le système des langues sémitiques et celui des langues indo-européennes, mais ces deux systèmes, ainsi que nous l'avons dit plus haut, comprennent la flexion d'une manière toute différente. Leurs racines sont tout à fait distinctes, leurs éléments formatifs sont essentiellement divers, et il n'y a nul rapport entre les deux modes de fonctionnement de ces éléments L'abîme n'est pas seulement profond entre les deux systèmes ; il est infranchissable.

« Quand deux langues peuvent-elles être scientifiquement tenues, dit H. Chavée, pour deux créations radicalement séparées ? Premièrement : quand leurs mots

(1) *Reise der œsterr. Fregate Novara um die Erde.* Linguistischer Theil, p. 273. Vienne, 1868.

simples ou irréductibles à des formes antérieures n'offrent absolument rien de commun, soit dans leurs étoffes sonores, soit dans leur constitution syllabique. Secondement : quand les lois qui président aux premières combinaisons de ces mots simples diffèrent absolument dans les deux systèmes comparés (1). »

C'est le cas des langues sémitiques et des langues indo-européennes, c'est le cas d'un nombre considérable de systèmes linguistiques. La conséquence de ce fait est grande.

Si c'est la faculté de langage articulé qui est la propre et la seule caractéristique de l'homme, ainsi que nous l'avons dit dans notre chapitre second, et si les différents systèmes linguistiques que nous connaissons sont irréductibles, ils ont pris naissance isolément en des régions bien distinctes. Il en résulte que les premiers êtres qui furent en voie d'acquérir la faculté du langage articulé ont gagné cette faculté en différents lieux à la fois et ont donné naissance ainsi à plusieurs races humaines originellement distinctes.

« Les anthropologistes français, dit le général Faidherbe (2), étaient généralement convenus que, la parole articulée distinguant seule radicalement l'homme des animaux, les précurseurs de l'homme ne devaient pas être désignés par le nom d'hommes, lorsqu'ils ne possédaient pas encore cet attribut. On comprend que ce n'est là qu'une affaire de mots, de convention. La seule chose importante, c'est de savoir si, chez cet être, qu'on l'appelle homme ou non, le langage a pris naissance sur un seul point, en une seule fois, ou bien d'une manière multiple, sous le rapport des lieux et des temps. Or, l'irréductibilité des langues humaines à une seule souche

(1) *Les Langues et les Races*, p. 13. Paris, 1862.
(2) *Essai sur la langue poule*. Paris, 1875.

prouve que la seconde hypothèse est la vraie. Si l'homme n'eût acquis cette faculté, conséquence des progrès de son organisation, que d'une manière unique, le langage fût resté sensiblement le même dans sa descendance, ou du moins on trouverait dans toutes les langues des traces de cette origine commune. La diversité extrême des langues et de leurs procédés prouve qu'elles ont été créées indépendamment les unes des autres, et probablement à des époques très différentes. Comme, en outre, les principales familles irréductibles de langues correspondent d'une manière générale aux grandes races de l'humanité, nous admettons que le langage a pris naissance d'une manière indépendante chez diverses variétés distinctes de ce que Frédéric Müller appelle l'*homo primigenius*, de ce que les anthropologistes français appellent les précurseurs de l'homme. »

La linguistique apporte ainsi aux polygénistes un argument capital. Leur arsenal était déjà bien fourni, il s'enrichit d'une arme nouvelle.

La langue étant un produit de la nature elle-même, étant la fonction d'un nouvel organe, il est évident que deux systèmes linguistiques irréductibles entre eux indiquent deux organes producteurs différents.

L'histoire nous apprend qu'un grand nombre de familles linguistiques se sont éteintes sans postérité ; cela est le fait de la concurrence vitale qui s'applique à la nature entière, partout et toujours. Plus nous remontons dans le cours des âges, plus nous trouvons de familles linguistiques indépendantes. C'est également le fait des races humaines.

Nous pouvons soutenir sans témérité que le primate précurseur de l'homme a dû acquérir sur bien des points à la fois ou successivement la faculté du langage articulé qui devait l'élever à la condition d'homme.

La linguistique nous conduit, en effet, à ce résultat,

en nous enseignant la multiplicité des systèmes linguistiques irréductibles (1).

§ 3. Dans la vie historique, les langues peuvent ne plus correspondre aux races.

On comprend ainsi, nous l'avons dit déjà, que dans la période historique de l'humanité il ne puisse plus naître de nouveaux systèmes de langues. L'origine du langage, l'acquisition de la faculté du langage articulé, étant identique avec la formation des premières races humaines, il s'ensuit que, le précurseur de l'homme une fois éteint, la formation de nouvelles familles linguistiques est absolument impossible. Tout effet nécessite une cause, et, la cause disparaissant, il n'y a plus d'effet possible.

Mais, lorsqu'elles sont entrées dans la période historique, les langues peuvent disparaître, comme peuvent disparaître les races.

Une foule d'exemples se présentent ici, parmi lesquels nous n'avons que l'embarras du choix.

On sait que les Finnois et les Lapons appartiennent à deux races essentiellement différentes ; pourtant le suomi, que l'on parle en Finlande, et la langue des Lapons, font partie d'une seule et même famille.

En Asie, nous voyons les différentes langues du groupe hindou acceptées par des populations se rattachant à une ou plusieurs races parfaitement distinctes de la race qui leur apporta leur système linguistique.

En Asie encore, et en Afrique, l'arabe est la langue courante, la langue maternelle d'un grand nombre de peuples qui ne font point partie de la race sémitique.

En Afrique, le système bantou est parlé à l'est par

(1) Sayce, Traduct. franc., p. 79, 95, 103.

des Cafres (zoulou, kafir), à l'ouest, dans la Guinée méridionale, par de véritables nègres africains ; or ces dernier et les Cafres ne doivent être confondus en aucune façon.

En Océanie, les Papous ont adopté, dans un assez grand nombre d'îles, des langues qui appartiennent à la famille maléo-polynésienne.

Les Malais, les Polynésiens appartiennent à deux types ethniques fort divers : leurs langues ont un fonds commun.

La sélection naturelle a fait disparaître dans le cours de l'histoire un nombre considérable d'idiomes : les langues qui se trouvent en collision nous offrent le spectacle des groupes animaux qui ont à lutter les uns contre les autres pour assurer leur existence. Il faut gagner sur ses concurrents ou se résigner à disparaître devant leurs progrès. De même que, dans le combat pour la vie et le développement, les races les mieux armées l'emportent finalement sur celles qui sont le moins favorisées, de même les langues qui sont le mieux servies par leurs propres aptitudes et par les circonstances extérieures l'emportent sur celles dont la force évolutive est moins considérable et sur celles que les conditions historiques ont moins bien préparées au combat. Sur notre territoire, nous voyons le français, la vieille langue d'oïl, avoir raison petit à petit des dialectes de langue d'oc, ses frères, de son autre frère le dialecte italien corse; de deux parents plus éloignés, le breton et le flamand ; d'un étranger, le basque. Dans les îles-Britanniques, l'anglais fait disparaître les langues celtiques : l'irlandais, l'écossais, le mannois, même le gallois ; il y a peu de temps, il en a terminé définitivement avec le cornique. L'allemand a eu raison d'un certain nombre d'idiomes slaves ; au moyen âge on a parlé slave jusque dans le Mecklembourg, le Brandebourg, jusque dans une

grande partie de la Saxe, de l'Autriche proprement dite et de la Carinthie. L'allemand moderne a étouffé un frère du lithuanien, le vieux prussien ; une langue slave, le polabe. Chaque jour il gagne sur deux autres langues slaves : le polonais et le vinde (ou sorbe de Lusace). L'espagnol est en voie d'en finir avec le basque, l'anglais avec les langues de l'Amérique septentrionale. Les Normands perdirent en France leur idiome scandinave, les Burgondes y perdirent également leur idiome d'origine germanique, comme les Lombards en Italie. En Italie déjà, le latin avait étouffé ses frères, l'osque et l'ombrien.

D'autres idiomes ont tenté de s'imposer violemment, mais n'ont pu réussir à se faire accepter. En Europe, par exemple, deux langues ouralo-altaïques se trouvent dans ce cas. L'une de ces deux langues est le turc. C'est en vain qu'il a été porté au cœur même de l'Europe ; il n'occupe plus aujourd'hui qu'une très minime partie de la Turquie européenne. Dans l'île de Candie les Turcs en sont arrivés presque tous à parler grec. La seconde langue ouralo-altaïque qui ait tenté de s'imposer à l'Europe, et dont la décadence actuelle ne saurait être sérieusement contestée, est le magyar. Ce n'est point qu'elle ne jouisse de privilèges considérables et que l'appui officiel ne lui soit acquis aux dépens des langues avoisinantes (1) ; on peut prévoir cependant que la langue des Magyars disparaîtra dans un avenir plus ou moins prochain.

Une partie des Hottentots abandonnent leur langue pour le hollandais. Dans l'Asie centrale bien des populations ouralo-altaïques ont adopté la langue persane. Il serait fastidieux de prolonger cette énumération.

(1) *Les Serbes de Hongrie*, p. 310. (Anonyme.) Prague, 1873.

§ 4. La transformation des espèces en linguistique.

Entrées dans la vie historique, les langues ne tardent pas à voir leur système phonétique s'altérer et leurs formes se modifier petit à petit. Les consonnes et les voyelles se transforment souvent en consonnes plus fortes ou plus faibles, en voyelles plus aiguës ou plus profondes ; les unes et les autres exercent souvent une influence réciproque qui s'accuse de plus en plus, et les diverses branches d'une seule et même famille, ayant chacune leurs principes particuliers de modification, s'éloignent chaque jour un peu plus les unes des autres.

Le persan et le français sont bien plus distants l'un de l'autre que ne l'étaient le vieux perse et le latin; l'anglais et l'allemand sont séparés l'un de l'autre par un intervalle bien plus considérable que celui qui séparait l'anglo-saxon de l'ancien haut-allemand.

Et non seulement les formes s'altèrent et se modifient, mais parfois aussi elles se perdent totalement; la langue commune indo-européenne possédait huit cas, le latin n'en connaissait guère plus que les deux tiers, il n'en restait que deux dans la langue d'oïl du moyen âge, et le français moderne les a totalement perdus; la langue sémitique commune en possédait trois : seul de tous ses rejetons l'arabe littéral les a conservés.

A vrai dire, cela n'est pas une transformation, c'est une dégradation.

La transformation vraie, celle dont nous avons à parler en ce moment, c'est la variation de l'espèce, Il y a longtemps déjà que la variabilité de l'espèce, en linguistique, est un fait acquis à la science, et que ceux-là seuls pourraient encore révoquer en doute qui prennent l'étymologie pour la linguistique.

Les phases de cette évolution, telles que nous les saisissons actuellement, sont celles de la formation, de la croissance, de la plénitude, de la décadence. La variation est continuelle; les langues naissent, se développent, entrent en décadence, s'éteignent comme tous les êtres organisés. Leur développement historique se modifie dans le cours des âges suivant telles ou telles conditions, cela est incontestable, mais l'observateur de ces modifications ne saisit jamais en elles que des phénomènes d'évolution naturelle : la preuve évidente de ce fait, c'est que l'évolution est sommairement la même dans des familles linguistiques essentiellement différentes les unes des autres.

Dans ses *Recherches sur les langues tartares*, Abel Rémusat a bien indiqué la nature de l'évolution générale des idiomes : « En les étudiant avec attention, dit-il, on est tenté de croire qu'ils sont aussi constants dans leur marche que la constitution physique qui leur a donné naissance... Peut-être règne-t-il dans les langues moins d'arbitraire qu'on n'a coutume de le supposer; et si l'on y portait le scrupule nécessaire, peut-être trouverait-on à y prendre des signes aussi sûrs, aussi prononcés, aussi caractéristiques que ceux qu'on peut tirer de la physionomie, de la couleur de la peau ou de celle des cheveux, ou de toute autre particularité physique et extérieure. » Ce « scrupule nécessaire » a été porté : nous allons voir à quelles conclusions il a conduit.

Nous ne connaissons aucune langue dans son état embryonnaire, s'il est permis de s'exprimer ainsi; toutes les langues soumises à cette observation directe, même celles des populations qui se trouvent aux derniers ou, pour mieux dire, aux premiers degrés de l'échelle humaine, ont passé la période de formation, qui a été préhistorique, et sont actuellement dans la période historique, généralement dans leur décadence. Mais en sépa-

rant méthodiquement et en comparant leurs éléments formatifs, on peut se rendre compte de ce qu'a été l'ancienne période de formation.

Le résultat de ces recherches comparatives a confirmé la théorie formulée en 1818 par Guillaume Schlegel : les langues ont tout d'abord passé par une période monosyllabique; un grand nombre se sont élevées à la phase de développement dite *phase agglutinante*, et, parmi ces dernières, quelques-unes enfin, le plus petit nombre, ont atteint une dernière phase, celle de la flexion. La structure des premières est simple, la structure des secondes est complexe, la structure des dernières est plus complexe encore.

Dans la première phase, le mot et la racine sont tout un, et chaque mot-racine, chaque racine-mot, pour mieux dire, est monosyllabique. La phrase est dès lors une pure et simple succession de racines isolées. Il est de toute évidence que tel a été le premier procédé d'élocution : on s'exprimait en mettant à la suite les uns des autres des monosyllabes, qui devaient parfois, on n'en saurait douter, être des onomatopées, des imitations de bruits, de sons, de cris.

Les langues monosyllabiques actuelles ont singulièrement amélioré ce procédé très primitf, et elles l'ont fait tout en restant monosyllabiques. Elles n'ont point créé de grammaire, ne connaissant point de structure dans les mots, mais elles ont créé une syntaxe. Cette syntaxe consiste dans la position donnée dans la phrase aux différentes racines-mots; la place qu'occupe le monosyllabe dans l'ensemble de la phrase précise le rôle du monosyllabe en question. Ce procédé d'ordre tout syntactique revient forcément en usage dans les langues analytiques actuelles, qui sont le plus avancées en décadence. Lorsque, par exemple, nous disons en français : « Pierre aime Jean », il est de toute nécessité que nous

placions le mot *Pierre* en tête de la phrase, le mot *Jean* à la fin, car ces deux mots ont perdu toute la distinction morphologique qui pouvait faire de chacun d'eux soit un sujet soit un régime : la place qu'occupe le mot *Pierre* dans la phrase en question indique qu'il est sujet, la place qu'occupe le mot *Jean* indique qu'il est régime. Dans les langues synthétiques (dont il sera parlé tout à l'heure), il n'en est pas ainsi ; le sujet, le régime sont distingués par leur forme même, et la position dans la phrase est sans aucune importance ; on dit indifféremment « Helvetii legatos miserunt », ou « legatos miserunt Helvetii », les deux noms révélant leur fonction par leur forme même.

En chinois, par exemple, la racine qui dans une phase doit valoir comme sujet, comme nominatif, se place avant la racine qui doit affecter le sens verbal ; en assignant ainsi au mot qui doit être sujet une place fixe dans la phrase, on obvie au manque d'éléments grammaticaux qui, en latin, en grec, caractérisent le cas nominatif, par exemple l'*s* de *dominus*, de *logos*. Dans une langue monosyllabique, en somme, point de grammaire ; point de formes nominales, point de formes verbales, ni déclinaison. ni conjugaison, point de genre, point de modes ni de temps, rien qu'une syntaxe. C'est d'ailleurs ce que l'on saisira plus aisément en étudiant la transition du monosyllabisme à l'agglutination, le passage de la première à la seconde phase linguistique.

Cette transition, cette évolution s'opéra d'une façon tout à fait simple. Certains mots-racines abdiquèrent une partie de leur sens, devinrent de simples éléments de relation, de rapport, tandis que les autres mots-racines conservèrent leur sens dans toute sa plénitude, dans toute son indépendance. En chinois, et dans les autres langues monosyllabiques actuelles, on trouve cette division des mots en mots « pleins » (que nous pouvons en

français traduire par un verbe, par un nom) et en mots « vides » dont le sens primitif s'est peu à peu obscurci et qui, peu à peu ont servi à déterminer, à préciser la notion large des mots pleins. Ce procédé a été employé beaucoup plus tard, chose fort intéressante, par des langues arrivées à un haut degré de développement. En latin, par exemple, à côté du mot *circus*, cercle, se trouve le mot *circum*, qui veut dire « autour » ; or ce dernier n'est plus qu'une espèce de mot « vide », un mot n'indiquant que la relation : *qui circum illum sunt*, ceux qui sont autour de lui : *circum hæc loca*, autour de ces lieux, dans ces environs. De même, à côté de *vertere*, *verto*, se trouve *versus : versus ædem Quirini ;* de même, à côté de *tenuis*, étendu, délié, de *tensus*, tendu, se trouve *tenus : crurum tenus*. jusqu'aux jambes.

Ce que devait faire le latin, qui, du mot « plein » *circus*, cercle, a tiré le mot « vide » *circum*, autour (le premier conservant son sens intégral, le second ne devenant plus qu'un élément de relation), cela même, les langues monosyllabiques l'ont fait pour arriver à plus de clarté dans l'expression. Ainsi le mot employé en chinois pour signifier « avec » et qui rend le cas instrumental (avec le bras, avec un bâton), est simplement la racine qui, étant « pleine », signifiait « se servir de, faire usage de ».

Dans les langues monosyllabiques, les mots pleins et les mots vides se suivent, sont mêlés les uns aux autres sans jamais s'amalgamer ; en autres termes, les racines sont toujours isolées les unes des autres, il n'y a pas de mots comprenant plusieurs syllabes. A la vérité, on peut former des sortes de composés en rapprochant (sans toutefois les souder) deux mots différents : ainsi, en chinois, le mot *fú*, *père*, et le mot *mú*, *mère*, rapprochés sous la forme *fú-mù*, donnent le mot « parents »; rapprochés de même, les mots signifiant « loin » et

« près » donnent le mot signifiant « distance ». Mais il n'y a encore ici aucune dérivation; des deux racines, aucune ne sert à l'autre d'élément de relation, chacune garde toute sa personnalité.

A un moment donné du développement linguistique, un pas de plus est fait : le mot indiquant la relation, le mot « vide » s'accole au mot « plein » et une forme polysyllabique, une forme agglomérante, prend naissance. Le mot est dorénavant formé autrement que par une simple racine isolée, il consiste en éléments divers agglomérés : nous en sommes à la période morphologique secondaire, à la période dite *d'agglutination* ou *d'agglomération*. Qu'on le remarque bien d'ailleurs, il ne s'agit pas ici de deux mots « pleins » se réunissant pour former un composé : il s'agit — ce qui est bien différent — de l'agglomération au mot principal d'un mot jouant le rôle secondaire d'élément dérivatif, précisant les relations de la racine principale à laquelle il se soude. Cet élément secondaire, ce dérivatif, est par exemple *ta* dans les mots sanskrits *çruta*, entendu; *mata*, pensé ; *bhrta*, porté ; jadis, cet élément dérivatif *ta* a eu son indépendance, il a eu un sens « plein », il n'est plus arrivé, par la suite des temps, qu'à servir d'élément dérivatif, d'élément de relation.

Lorsque cet élément dérivatif est placé après la forme radicale, il est appelé *suffixe* (*ter* est suffixe dans *pater*, *mater*, *frater*); lorsqu'il est placé en tête du mot, il est appelé *préfixe :* c'est le cas chez les Cafres; parfois, l'élément dérivatif est intercalé dans le corps même de la racine, et il prend alors le nom d'*infixe :* ce mode de dérivation est rare.

Il est bon d'ajouter que la dérivation n'a point de limites, qu'un mot dérivé peut être dérivé à son tour, ce dernier également, et ainsi de suite. Ainsi, en magyar, le dérivé *zárat* signifie « il fait fermer » ; le dérivé

zárhat « il peut fermer » au moyen d'une dérivation secondaire, on forme *zárathat* « il peut faire fermer » ; *záratgat* « il fait fermer souvent » est également secondaire et *záratgathat* « il peut faire fermer souvent » est un dérivé tertiaire. Les langues de la troisième période d'évolution, par exemple le latin, présentent un nombre considérable de dérivés de cette sorte, dérivés primaires, secondaires, tertiaires, etc. Le mot *pater* est un dérivé primaire dont l'élément plein ou radical est *pa*, et dont l'élément dérivatif est *ter ; paternus* est un dérivé secondaire; on a ensuite *paternitas*, etc. En tout cas, nos langues n'ont pas l'extraordinaire faculté de dérivation que possèdent un certain nombre d'idiomes simplement agglutinants. « Tant de choses en deux mots ? » dit le Bourgeois gentilhomme, et Covieille lui répond : « Oui. La langue turque est comme cela, elle dit beaucoup en peu de paroles. » Ce qu'il y a d'exact, c'est qu'en un seul et même mot la langue turque peut introduire un certain nombre de notions : elle dit en un seul mot non seulement *sêvmèk*, aimer, mais encore *sèvmèmek*, ne pas aimer ; *sèvilmèk*, être aimé ; *sèvilmèmèk*, ne pas être aimé ; *sèvdirmèk*, faire aimer ; *sèvdirmèmèk*, ne pas faire aimer ; *sèvinmèk*, s'aimer, et ainsi de suite : les éléments dérivatifs indiquent, dans ces diverses formes, la négation, l'idée de cause, l'idée de retour sur soi-même, autant de notions que le français doit exprimer par plusieurs mots.

La plus grande partie des langues en est à la seconde période morphologique, à la période agglutinative, par exemple les langues des nègres occidentaux et orientaux, celles des Malais, des Polynésiens, des Dravidiens, des peuples altaïques, le basque, les langues américaines, etc., etc. Mais la communauté de structure ne préjuge point la parenté ; le fait de ce que deux langues en sont à la même phase d'évolution n'accuse en rien une

communauté d'origine, et il en est de même, bien entendu, en ce qui concerne les langues monosyllabiques.

Il y a lieu, d'autre part, de ne pas négliger ce fait, que dans la phase d'agglutination, certaines langues ont peu progressé, que d'autres, au contraire, ont singulièrement avancé. Certaines langues des nègres de l'Afrique occidentale usent encore, à côté de formes agglutinées, de procédés propres au monosyllabisme : il n'y a point là retour à d'anciennes formes, mais bien maintien d'anciennes formes au milieu de formations plus complexes. Il faut enfin ajouter que les formes de certains idiomes trahissent perpétuellement le passage du monosyllabisme à l'agglutination. On peut citer le khassia, parlé au nord-est de l'Inde par 200.000 individus environ. Certes, cette langue n'a point de valeur littéraire, elle appartient à un peuple qui ne connaît qu'une civilisation très rudimentaire; mais pour celui qui étudie les phénomènes de l'évolution linguistique, elle a une importance de premier ordre, et l'on en pourrait dire autant de bien d'autres idiomes dont le philologue n'a souci. C'est ainsi que d'obscures espèces végétales ou animales sont souvent plus riches d'enseignements pour le botaniste et le zoologue que ne le sont tant d'espèces communément recherchées pour leur utilité pratique ou même pour leur simple beauté.

S'il est aisé de donner l'explication du passage de la première phase linguistique à la seconde, il est, par contre, beaucoup moins facile, nous l'avons dit, d'exposer le phénomène d'évolution de l'agglutination à la flexion. En principe, cette évolution a lieu par le fait d'une modification phonique de la racine : par exemple, en arabe, *katab* il a écrit, *katib* écrivant, *me-ktub* écrit. En ce qui concerne la flexion du système indo-européen (sanskrit, perse, grec, latin, etc.), l'évolution aurait eu lieu, d'après Victor Henry, non seulement au moyen de modifications

phoniques de la racine, au moyen d'une modification dans l'accentuation, mais encore au moyen d'une agglutination par infixes; dans son tableau systématique des racines indo-européennes, Chavée, qui a rendu, il y a quarante ans, à la science du langage des services que l'on ne saurait oublier, Chavée a été sur la voie de cette interprétation.

Quoi qu'il en soit donc de ce point encore assez obscur du passage du second au troisième état de structure linguistique, si nous considérons les anciennes langues indo-européennes telles que le sanskrit, le grec, le latin, nous reconnaissons qu'elles sont, à divers degrés, synthétiques, et si nous recherchons la nature des langues romanes (français, espagnol, italien, etc.), nous trouvons qu'elles sont analytiques. Telle est, en effet, l'œuvre de la décadence linguistique, décadence moins hâtive dans les langues slaves qu'en allemand, moins hâtive en allemand que dans les langues romanes.

Cette décadence, qui constitue une nouvelle phase d'évolution, ne se produit pas au hasard.

Si nous envisageons tout d'abord la phonétique, nous avons à constater les résultats du moindre effort : les diphtongues se condensent, *veicos*, *deivos*, deviennent *vîcus*, *divus ;* l'assimilation s'exerce aussi bien sur les consonnes que sur les voyelles : *notte*, *sette*, *atto*, répondent à *noctem*, *septem*, *actum ;* en grec, une aspiration (l'esprit rude) répond à une sifflante primitive (ἑπτά, *septem*, ἑκυρός, *socer ;* ἕρπω, sanskrit *sarpâmi ;* ἕδος, sanskrit *sadas*). Un grand nombre de variations phoniques qui déroutent à première vue quiconque est peu familiarisé avec les études linguistiques, se justifient par le rapprochement avec d'autres mots. Que le mot français *sache* vienne du latin *sapiam*, cela tout d'abord paraît étrange; mais l'étonnement tombe lorsque l'on voit *sepia* donner *sèche*, *Clipiacum* donner *Clichy*, *apium* donner *ache*.

Le fait s'explique encore mieux lorsqu'en face du latin *pi* devenant *ch* on voit le latin *bi* devenant *j* (*g* doux) : c'est le cas pour les mots *rage* et *rouge*. Il y a là un parallèle phonique qui montre à quel point les variations de cette sorte dépendent d'une évolution naturelle.

C'est sur le fait du moindre effort que repose la formation phonique de l'ancienne langue d'oïl, du français. Le français maintient la syllabe latine qui porte l'accent et sacrifie les syllabes suivantes : *tabula*, *fémina*, *régula* deviennent *table*, *femme*, *règle*. (Déjà, dans les textes latins, on trouve *vinclum*, *poclum*, *periclum*, *oraclum*, *vehiclum*.) Les mots français *porche*=*pórticus*, *frêle*=*fragilis*, *roide*=*rigidus*, sont des mots naturellement et régulièrement formés : leurs doublets *portique*, *fragile*, *rigide* sont des formations relativement récentes, dites *savantes*, en réalité des calques barbares. Parfois la recherche du moindre effort fait introduire dans les mots une consonne adjuvante : *b* dans *humble*, *comble*, *sembler*, *nombre* (de *humilis*, *cumulus*, *simulare*, *numerus*), *d* dans *pondre*, *tendre*, *gendre* (de *pónere*, *téner*, *géner*). Parfois des composés se contractent : *magis volo*, je préfère, devient *malo*, *potis esse* devient *posse* ; *lapicida* tailleur de pierres, est pour *lapidicida*, *cordolium*, crève-cœur, chagrin, est pour *cordidolium* ; le latin *idololatres*, tiré du grec, a donné naissance à une forme *idololâtre*, qui s'est condensée en idolâtre ; l'anglais *lord* répond à un *lauard* plus ancien, qui lui-même est pour *hlâf weard*, dispensateur de pain.

La décadence, en ce qui concerne la grammaire, répond de même à une simplification. L'ancienne langue indo-européenne, que la comparaison du sanskrit, du latin, du grec, des langues slaves, des langues germaniques a permis de restituer dans ses formes importantes, possédait une riche déclinaison. Le latin a perdu une partie des cas de cette déclinaison, ne possède de

tels autres cas que des vestiges (*humi*, à terre ; *belli*, en temps de guerre ; *domi*, à la maison) ; l'ancien français fait un pas de plus, ne conserve plus que deux cas : un cas sujet et un cas régime (aussi bien direct qu'indirect). Au quatorzième siècle, cette déclinaison très simplifiée disparaît de la langue française devient complètement analytique. Ce n'est pas sans avoir gardé des traces de la déclinaison du moyen âge : *pâtre* est l'ancien cas sujet (*pastre*) répondant au nominatif latin *pástor ; pasteur* est l'ancien cas régime répondant au latin *pastórem ; sire* est l'ancien cas sujet ; *seigneur*, l'ancien cas régime ; il en est de même de *chantre* et *chanteur*. En principe, c'est le cas sujet du français du moyen âge qu'a laissé tomber le français moderne, c'est le cas régime qu'il a conservé, le faisant servir tout à la fois de sujet et de régime. Ainsi le nominatif *latro* donnait *li lerres ;* l'accusatif *latrónem* donnait *le larron ;* le nominatif *ábbas* donnait *li abes ;* l'accusatif *abbótem* donnait *le abbé*. C'est la forme accusative qui a seule persisté, servant à la fois et pour le régime et pour le sujet. Parfois, cependant, c'est le cas sujet qui a résisté, comme le prouve la consonne *s* de *fils* (sujet *li fils*, régime *le fil*, latin *filius*, *filium ;* de *bras* (sujet *li bras*, régime *le brac*).

La simplification de la déclinaison se retrouve dans toutes les langues modernes. En persan, il n'y a plus, à proprement parler, de déclinaison : lorsqu'on veut exprimer le datif, l'accusatif, on joint au nom certaines prépositions ; on rend le génitif par un procédé syntactique. Le grec moderne a perdu les formes du duel et le datif. Si nous considérons les langues sémitiques, nous voyons que l'arabe courant, l'arabe parlé, laisse tomber les désinences qui, dans l'arabe littéraire, indiquent les trois cas du système linguistique sémitique ; dans l'arabe vulgaire, ces cas se reconnaissent par la position des mots dans la phrase ou par l'emploi de prépositions.

Passant à la conjugaison, nous rencontrons les mêmes phénomènes d'analytisme. En voici un ou deux exemples : dans le système indo-européen, le parfait était formé par le redoublement de la racine (λέλοιπα, *cecini*). Le latin forme déjà des parfaits au moyen de compositions de mots : *amavi, audivi,* où *vi* est pour *fui,* comme le prouve les formes ombriennes en *fei.* L'imparfait *amabam,* le futur *ama-bo* sont également composés. Le français va plus loin et donne les formes analytiques : *j'ai aimé, j'avais aimé.* Le futur *j'aimer-ai est pour* « j'ai à aimer » ; c'est ce que confirment les vieilles formes méridionales : *dar vos n'ai,* je vous en donnerai, *dir vos ai,* je vous dirai.

La décadence linguistique provient parfois de ce que la valeur primitive d'une forme, d'un mot, a été oubliée. Les formes latines qui ont donné naissance aux mots *luette, lierre,* étaient *uveta, hedera :* en ancien français ces mots latins étaient devenus *uette, hierre.* On disait avec l'article *l'uette, l'hierre :* la méconnaissance de la valeur et du rôle de l'article l'a fait annexer aux mots en question et l'on dit *la luette, le lierre.* Certains patois ont conservé la forme ancienne et disent encore *hierre.*

A la sélection qui s'applique aux différents idiomes d'une même famille, ou à des familles distinctes les unes des autres, il y aurait lieu d'ajouter la sélection qui s'applique dans un seul et même idiome, soit à l'usage de telles ou telles formes, soit à l'usage de tels ou tels mots. C'est ici que l'étude des patois est d'un précieux intérêt. Les patois ne doivent pas être regardés comme des dégénérescences des langues littéraires : les langues littéraires sont des dialectes heureux ; les patois sont, au contraire, des dialectes malheureux, des dialectes qui n'ont point passé à la condition de langues littéraires. Mais à chaque instant, dans les patois, nous rencontrons des formes, des mots, que les langues litté-

raires, leurs sœurs, n'ont point conservés. De là l'importance considérable des patois dans l'histoire naturelle du langage. Et il ne faudrait pas croire que ces rencontres, dans les patois, de vieilles formes, devenues inconnues aux langues littéraires, soient exceptionnelles.

En nous en tenant simplement à la langue littéraire elle-même, combien de mots voyons-nous subsister n'ayant plus qu'un emploi très particulier et très précis, qui jadis avaient une acception générale et courante. Le latin *cogitare*, penser, a donné à l'ancien français *cuider :* ce mot a disparu de la langue littéraire, mais un témoin nous en est resté dans « l'outrecuidance ». Le latin *faber*, artisan, fabricant, a donné à l'ancien français *fevre* (*li fevres*, l'ouvrier), que nous retrouvons comme composant dans « orfèvre ». Le latin *fons*, *fontis*, fontaine, ne se retrouve que dans la locution de « fonts baptismaux ». En dehors de ces emplois particuliers, les formes *cuider*, *fevre*, *fonts* sont, dans la langue française littéraire, comme des formes fossiles (Cf. Ars. Darmesteter, *la Vie des mots*, Paris, 1887).

La perte d'un grand nombre d'idiomes a eu ceci de fâcheux pour le progrès des études linguistiques, que ç'a été souvent la disparition d'autant de formes intermédiaires dont l'existence eût expliqué une foule de formes actuellement vivantes. En cela encore, ce qui se présente dans les langues est tout à fait comparable à ce qui se passe dans la vie des espèces végétales ou animales. Ajoutons qu'une espèce linguistique une fois éteinte, aucune circonstance ne peut la faire revivre. Il y a peu de temps qu'ont succombé les Tasmaniens et que leur langue a disparu avec eux : pas plus qu'ils ne pourront reparaître, eux qui avaient été le produit d'une longue évolution ethnique, pas plus ne pourra reparaître un langage semblable au leur, qui avait été, lui aussi,

le produit d'un long développement. C'est ainsi que dans le monde végétal et le monde animal la disparition d'une espèce est toujours définitive : pour l'amener à une vie nouvelle, il faudrait le retour impossible des conditions de toutes sortes qui l'avaient amenée à l'état qu'elle présentait au moment de sa disparition.

On a souvent objecté aux partisans de la doctrine du transformisme l'absence d'intermédiaires entre les formes actuellement existantes et les formes plus anciennes. Nous n'avons pas à nous prononcer ici sur une question de zoologie ou de botanique, mais nous devons faire remarquer que lorsqu'il s'agit du langage, cette objection ne peut même pas être posée. Ici le procédé d'évolution est facile à suivre, on le prend en voie d'exécution. On voit comment un idiome à flexion a dû passer par la forme de l'agglutination et comment toute forme agglutinante a dû passer auparavant par la forme du monosyllabisme. La transformation de l'espèce est ici un fait patent, et nous pouvons dire que cette transformation est l'un des principes fondamentaux de la science du langage (1).

N'est-ce pas une preuve nouvelle et bien éclatante de ce fait dont nous avons eu à nous occuper tout au commencement de ce livre, que la linguistique est avant tout une science naturelle ?

Un mot encore avant de terminer.

Nous avons parlé tour à tour de pluralité originelle et de transformation. Ces deux termes, aux yeux de quelques personnes, sembleraient peut-être se contredire ; en fait il n'en est rien et ils se concilient sans difficulté.

La doctrine de la pluralité originelle des langues et des races humaines n'a pas la prétention de faire échec

(1) Withney, *Language and the Study of Language*, 3e édit., p. 175. Londres, 1870. — *La vie du langage*, trad. franç., 1875. — Powel, *the Evolution of Language* (Transact. of the Anthrop. Soc. of Washington, 1881). — Journal *l'Homme*, t. I, p. 609.

à la doctrine plus générale de l'unité cosmique. En fin de compte, il faut reconnaître toujours que toutes les formes existantes, toutes sans exception, ne sont que les différents aspects de la matière, qui est *une* comme elle est infinie. Mais cette unité n'empêche en aucune façon que telles ou telles formes identiques, analogues si l'on veut, ne se soient développées simultanément en des centres différents.

En définitive, l'usage du langage et de la vie du langage est en présence de développements successifs d'ordre tout à fait naturel. Le perfectionnement organique du cerveau dote le premier des primates de la faculté du langage articulé ; cette faculté, mise en jeu, donne naissance à un système très rudimentaire d'expression, ayant sa source, comme l'a dit fort justement Lucrèce, dans un besoin impérieux.

Le besoin est, en effet, le créateur des mots. Peu à peu, les monosyllabes se différencient en mots principaux et en mots de signification secondaire ; une nouvelle phase naît avec le rapprochement plus intime des mots et les divers procédés de dérivation se développent de plus en plus. La troisième phase est caractérisée d'abord par un synthétisme remarquable, mais qui ne tarde point à se simplifier : une marche plus rapide de la civilisation est sans nul doute la cause de cette évolution nouvelle : la précision analytique s'accentue de plus en plus. La dernière forme n'est point atteinte évidemment par les langues française et anglaise ; mais, de même que le langage est né avec l'homme, puisqu'il est la grande caractéristique de l'humanité — caractéristique lentement et laborieusement conquise — de même il n'aura été transformé en un mode plus parfait d'expression que le jour où celui qui est actuellement le premier des primates aura gagné dans l'échelle des êtres un échelon supérieur à celui qu'il occupe aujourd'hui.

TABLE ANALYTIQUE

Abyssinie. Langues sémitiques de l'—, 240.
Accadien. Double sens attaché à ce mot, 197.
Accent, L' — latin, 312. Rôle de l' — latin dans la formation des langues novo-latines, 322.
Afghan. Dialecte éranien, 294.
Agaou, Dialecte éthiopien, 256.
Agglutination. Seconde forme linguistique, 61.
Aïnos. Langue des —, 189.
Akoucha, 191.
Albanais. Langue indo-européenne non classée, 400.
Aléoutiens. Dialectes —, 189.
Aljourou, 99.
Algonquin. Grammaire de l'—, 184.
Allemand. Caractéristique de l'— moderne, 364. Orthographe de l'—, 365.
Américaines (*Langues*). Leur grand nombre, 173. Quel serait leur caractère commun, 176. Ne se distinguent pas des autres langues agglutinantes, *ibid*. Vocabulaire des —, 187.
Amharique. Parent du ghez, 242.
Anglais. Différentes périodes de l'—, 356.
Anglo-saxon, 355.
Annamite. Est une langue indépendante, 52. Est une langue monosyllabique, *ibid*.
Anthropoïdes. Primates arrêtés dans leur développement, 37.
Appalache, 175.
Arabe. Groupe — des langues sémitiques, 234. L'arabe proprement dit —, 235. Son alphabet, 235. Sa place dans l'ensemble des langues sémitiques, 236. L' — vulgaire, 239. Dialectes de l'— vulgaire, 239.
Araméen. Groupe — des langues sémitiques, 218.
Araucan, 176.
Arévaque, 175.
Arménien. Sa place dans le groupe des langues éraniennes, 286.
Armoricain. Breton proprement dit ou —, 343.
Arya. Valeur de ce mot, 268.
Aryaque. Nom donné par quelques auteurs à la langue commune indo-européenne, 269.
Aryen. Valeur de ce mot, 268.
Asie Mineure. Langues indo-européennes de l'—, 398. Le grec parlé sur les côtes de l'—, 302. Le turc parlé dans l'intérieur de l'—, 147.
Assyrien. Est une langue sémitique —, 221.
Athapasque, 174.
Australie. Langues de l'—, 95.
Avare, 191.
Avesta. Livre sacré du zoroastrisme, 281. Sa traduction en langue huzvârèche, 288.
Aztek, 175.
Badaga, 114.
Baga, 80.
Baghirmi, 83.
Baktrien. Nom donné à la langue zende, 283.
Bambara, 80.

Bantou. Groupe —, 83.
Bari, 83.
Bas-allemand. Groupe —, 354. Le bas-allemand proprement dit, ou plattdeutsch, 357.
Basque. Limites actuelles du —, 156. Le — recule surtout devant l'espagnol —, 158. A été longtemps étudié sans méthode, 158. Son état d'isolement, 161. Les plus anciennes traces du —, *ibid*. Ses nombreuses variétés, 162. Phonétique —, 163. Formation des mots en —, 165. Incorpore le régime direct, 168. N'est pas parent des langues américaines, *ibid*. Le vocabulaire —, 169. Origine du —, 170.
Battak, 99.
Bedja. Dialecte éthiopien, 256.
Béloutche. Langue éranienne, 294.
Berber. Nom général du libyen moderne, 253.
Biafada, 80.
Bicol, 98.
Birman. Est une langue monosyllabique, 55.
Bochimans. Langue des—,72.
Bola, 80.
Bondou. Dialecte poul. 91.
Bornou, 82.
Boughi, 100.
Boullom, 80.
Bouriate. Importance du — dans le groupe mongol, 150.
Brahoui. Difficile à classer, 125.
Breton. Rameau — du groupe celtique, 342, 344. Idiome — proprement dit, ou armoricain, 343.
Bulgare. Limites du —, 381. Mauvaise conservation de ses formes, 381.
Byzantin. Le grec —, 300.
Cafir, 85.
Cafres. Langues des —, 84.
Californien. Groupe des langues américaines, 175.
Caraïbe, 175.
Cas. Les trois — de la langue sémitique commune, 214. Les huit — de la langue indo-européenne commune, 260. Déclinaison à deux — de la langue d'oïl et de la langue d'oc, 324.
Catalan. Peut se rattacher à la langue d'oc, 330.
Caucase. Langues du — 190.
Celtique. Groupe —, 338. Les deux branches du groupe —, 340.
Celtomanie, 347.
Chaldéen. Sa place dans le groupe araméen, 218.
Chananéen. Groupe — des langues sémitiques, 225.
Cherbro, 80.
Chéroki, 175.
Chibcha, 176.
Chilouk, 83.
Chinois. Dialectes —, 43. Sa grammaire n'est que syntaxique, *ibid*. Système graphique du —, 46.
Congo, 85.
Copte. Procède de l'ancien égyptien, 251.
Coréen. Langue agglutinante peu connue, 110.
Cornique. Breton de la Cornouailles, 342.
Croate. Voir *Serbo-croate*.
Cunéiformes (*Inscriptions*), 194.
Dace. L'ancien — n'est pas encore définitivement classé, 397.
Dakota, 175.
Dankâli. Dialecte éthiopien, 256.
Danois. Sa place dans le groupe scandinave, 353.
Dayak, 100.
Dêvanâgari. Alphabet —,279.
Dikélé, 86.
Dinka, 83.
Dippil, 95.
Djagataïque. Turc —, 142.
Dongolavi, 93.
Doualla, 86.
Dravidien. Etendue du territoire —, 111. Langues dravidiennes, *ibid*. Leur ancienne extension, 113. Simplicité de la grammaire

dravidienne, 116. Pauvreté du vocabulaire —, 122.
Ecriture. — chinoise, 48. — annamite, 53. — siamoise, 54. — tibétaine, *ibid.* — japonaise, 107. — coréenne, 110. Ecritures diverses des langues maléo-polynésiennes, 105. Le tamoul possède une — particulière, 123. — sémitique, 216. — assyrienne, 217. — arabe, 235. — himyarite, 240. — égyptienne, 250. — dévanâgarî, 273. — zende, 282. — perse, 285. — arménienne, 288. — huzvârêche, 290. Ecritures italiques, 317. — slave, 367.
Egbé, 82.
Egyptien. Langue khamitique, 248. Sa grammaire, 249.
Ehkili. Parent de l'himyarite, 240.
Elbe. Slave de l'—, 382.
Elou ou *singhalais*, 277.
Eranien. Groupe — des langues indo-européennes, 279. Classification des langues éraniennes, 280.
Erse. Gaélique d'Ecosse, 341.
Esclavon liturgique. Nom du slave ecclésiastique, 368.
Escuara. Nom original du basque, 159.
Espagnol. Sa place parmi les langues, novo-latines, 333. Gagne sur le basque, 158.
Espèces. Transformation des — en linguistique, 425.
Esquimaux. Dialectes des —, 189.
Esthonien, 129. Grammaire de l'—, 132.
Ethiopien. Groupe — des langues khamitiques, 255.
Etrusque. Opinions diverses sur l'origine de l'—, 392. Appartient au groupe indo-européen, *ibid.*
Etymologie. Dangers de l'—, 16. Ce que doit être l'— 18.
Feloup, 80.
Fernando-Po. Langue de —, 86.
Filham, 80.
Finnois. Groupe — des langues ouralo-altaïques, 128.
Flamand, 357.
Flexion. Troisième forme linguistique, 204. En quoi elle consiste, *ibid.* — indo-européenne et — sémitique, 207.
Forme linguistique.— du monosyllabisme, 39. — de l'agglutination, 57. — de la flexion, 204.
Formose. Langue de —, 98.
Foutadjallo. Dialecte poul, 91.
Foutatoro. Dialecte poul, 91.
Français. Formation du —, 326. Double espèce de mots —, 324. Déclinaison à deux cas de l'ancien —, 326. Dialectes de l'ancien —, 328.
Frioulan. Ladin oriental, 333.
Frison. Branche du bas-allemand 358.
Ga, 82.
Gaélique. Branche — du groupe celtique, 338.
Galate. L'ancien —, 344.
Galicien. Parent du portugais, 334.
Galla. Dialecte éthiopien, 256.
Gallois. Breton du pays de Galles, 342.
Gâthâs. Dialecte zend des —, 282.
Gaulois. L'ancien —, 343.
Gbandi, 80.
Gbésé, 80.
Géorgien, 191.
Germanique. Groupe —, 347. Origine obscure de ce nom, *ibid.* Caractéristique du groupe —, 348. Place du gothique dans le groupe des langues —, 349.
Ghez. Appartient au groupe arabe méridional, 241.
Ghiliaks. Langue des —, 189.
Glagolitique. Ecriture —, 367.
Gônd. 114.
Gothique. La véritable orthographe de ce mot, 349. Place du — dans le groupe des langues germaniques, 350.
Grantha. Alphabet —, 123.
Grébo, 82.
Grec. Rameau — des langues

indo-européennes, 295. Ne forme par une branche particulière avec le latin, *ibid.* Sa grammaire, *ibid.* Ses dialectes, 298. La langue grecque commune, 299. Le — byzantin, *ibid.* Le — moderne, *ibid.* Extension du — moderne, 301. Prononciation du — ancien, 302.
Guanches. Langue des anciens —, 255.
Guarani, 175.
Guaycuru, 176.
Hadendoa. Langue des —, 256.
Haousa, 80.
Harari. Parent du ghez, 242.
Harmonie vocalique. L'— dans les langues ouralo-altaïques, 152.
Haut-allemand. Les trois périodes du —, 363. Deux sortes de —, 361.
Hébreu. Ses diverses périodes, 225. Sa grammaire, 226. Son alphabet, 230.
Héréro, 86.
Himyarite. Fait partie du groupe arabe méridional, 240.
Hindou. Groupe — des langues indo-européennes, 270. Langues néo-hindoues, 276. Phonétique des langues néo-hindoues, 277.
Hindoui. Son extension au moyen âge, 277.
Hollandais, 357.
Homme. L'— n'est — que par le langage, 22, 27, 38. Le précurseur de l'— et la linguistique, 37.
Hottentots. Langue des —, 68.
Huasteck, 175.
Huzvârèche. L'Avesta traduit en —, 288. Influence de l'araméen sur le —, 289. Sa grammaire, *ibid.* Son alphabet, 290.
Hyperboréennes (*Langues*). Ce nom général n'indique pas la parenté, 189.
Ibérienne (*Question*). La —, 170.
Ibo, 38.
Iénisséin. Samoyède —, 127.
Incorporation. L'— proprement dite diffère du polysynthétisme, 177. L'— en basque, 169. L'— dans les langues ouralo-altaïques, 139, 151.
Indo-européen. Dans le système — la flexion est autre que dans le système sémitique, 207. L'— et le sémitique sont irréductibles, 208. Langues —, 257. Langue commune —, 259. Comment se subdivisa la langue commune —, 269, 402. Où fut parlée la langue commune —, 407.
Indo-germanique. Dénomination vicieuse, 268.
Ingouche, 191.
Innuits. Dialectes —, 189.
Inscriptions cunéiformes. Langue de la seconde colonne des —, 194. — assyriennes, 221, — perses, 284.
Intonation. Importance de l'— dans les langues monosyllabiques, 47.
Iranien. Le nom d'éranien est préférable à celui d'—, 280.
Irlandais. Importance de l'— parmi les langues celtiques, grammaire de l'—, 340.
Iroquois, 174.
Islandais. Sa place dans le groupe scandinave, 353.
Isoubou, 86.
Italien. Sa place parmi les langues novo-latines, 331. Ses dialectes, 332.
Italique. Groupe. —, 306. Son degré de parenté avec le grec, 306. Alphabets italiques, 317.
Japonais. Rattaché sans raisons à d'autres langues agglutinantes, 105. Grammaire du —, 108.
Javanais, 99.
Kabyle. Dialecte lybien, 253.
Kalloum, 80.
Kamassin. Dialecte samoyède, 127.
Kamilaroi, 95.
Kamtchadal, 189.

Kanara, 112. Ecriture du —, 123.
Kanem, 82.
Kanori, 82.
Karaboukal, 191.
Karélien, 129.
Kasdo-scythique. Ce que quelques auteurs entendent par ce nom, 197.
Kasi-koumuque, 191.
Koyuga, 175.
Kénaï, 174.
Khamitique. Groupe —, 246. Nom défectueux, *ibid*. Hypothèse sur l'ancienne extension des langues —, *ibid*. Les langues — sont alliées aux langues sémitiques, *ibid*. Grammaire générale des langues —, 247. La branche — se distingue en trois groupes, 248.
Kihiaou, 85.
Kikamba, 85.
Kinika, 85.
Kinki. Dialecte australien, 95.
Kipokomo, 85.
Kirghizes. Langue des —, 143.
Kisouahili, 85.
Kissi, 80.
Kiste, 191.
Koinberri. Dialecte australien, 95.
Koldadji, 93.
Koloche, 175.
Kondjara, 93.
Koriaque, 189.
Kôta, 114.
Kotte, 189.
Koudagou, 112.
Koumalé, 93.
Kourde. Langue éranienne, 293.
Krévin. 129.
Kri, 174.
Krou, 82.
Kuéva, 175.
Kurine, 191.
Ladin. Les trois groupes du —, 332.
Lak ou *Kasi-koumuque*, 191.
Lamour ou *Ingouche*, 191.
Lamoute, 148.
Landoro, 80.
Langage articulé. La faculté du — est la caractéristique de l'homme, 22, 27, 38. Abolition de la faculté du —, 29. Localisation de la faculté du —, *ibid*. L'exercice de la faculté du — est un art, 35. Corrélation de la naissance de l'homme et de l'apparition de la faculté du —, 37. La caractéristique tirée de la faculté du — n'est que relative, 38.
Langue. La vie des —, 9. Il n'y a pas de — mixtes, 10. Les — monosyllabiques, 39. Différenciation des — agglutinantes d'avec les — monosyllabiques, 57. Pluralité originelle des —, 419. Comment se reconnaît la parenté des —, 413. Dans quelle condition la — peut ne plus correspondre à la race, 422.
Lapon, 133.
Latin. Relation de la parenté du — et du grec, 295, 306. Différence du vieux — et du — classique, 307. Phonétique du —, 308. Prononciation du — classique, 310. Accent — 312. Le — populaire donne naissance aux langues novo-latines, 319.
Laze, 192.
Lette. Limites du —, 392. Moins bien conservé que le lithuanien, *ibid*.
Lettique. Groupe —, 389. Ses dialectes, *ibid*. Est distinct du groupe slave, 394.
Libyen. Groupe — des langues khamitiques, 252.
Linguiste. Distingué du philologue, 8. N'est rien moins qu'un faiseur d'étymologies, 20.
Linguistique. Distinction de la — et de la philologie, 1. Ce que c'est que la —, 4. Aide que la — peut attendre de la philologie, 12. Argument que la — fournit aux polygénistes, 421.
Lithuanien. Son état de conservation, 390. Ses limites, *ibid*. Sa grammaire, 391.

Live. Dialecte finnois, 130.
Logoné, 83.
Lusace. Slave de —, 376.
Lycien. Langue indo-européenne de l'Asie Mineure, 398.
Maba, 83.
Madurais, 100.
Magyar. Son importance dans le groupe finnois, 134. Limites du —, 135. Sa grammaire, 137.
Makassar, 100.
Makoua, 85.
Malai. Groupe — des langues maléo-polynésiennes, 98.
Malayàla, 112. Ecriture du —, 123.
Maléo-polynésiennes (*Langues*). Classification des —, 97. Leur origine commune, *ibid*. Forment un système indépendant, *ibid*. Leur grammaire, 101.
Malgache, 98.
Mallais. Est d'origine arabe, 239.
Mame, 175.
Mandchou. Fait partie du groupe tongouse, 148.
Mandingue, 80.
Mannois. Gaélique de l'île de Man, 340.
Mano, 80.
Maya, 175.
Mélanisien. Groupe —, 97.
Mendé, 80.
Métamorphose. Période de — régressive des langues, 10.
Mikmak, 174
Mingrélien, 192.
Mitchi, 83.
Mixtek, 175.
Mohican, 174.
Mongol. Groupe — des langues ouralo-altaïques, 150.
Monosyllabiques (*Langues*), 39. Ce qu'est la grammaire des —, 42.
Monosyllabisme. Première forme linguistique, 39.
Mordvin, 129, 133.
Morphologie. Ce que c'est que la —, 5. La — n'établit pas à elle seule la parenté des langues, *ibid*.
Mosarabe. Est d'origine arabe, 239.
Mosgou, 83.
Mourio, 82.
Mpongoué, 86.
Nahuatl ou *aztek*, 175.
Néerlandais, 357.
Nègres. Langues des — d'Afrique, 73.
Négritos. Langue des —, 94.
Néo-calédonien, 95, 97.
Néo-grammairiens. Théorie des — sur le système indo-européen, 266.
Ngourou, 82.
Nogaïque, 143.
Nordique Ancien —, 352.
Norvégien. Sa place dans le groupe scandinave, 353.
Nouer, 83.
Noupé, 83.
Novo-latines (*Langues*), 318. Formation des — 319. Eléments étrangers dans les —, 321. Les sept —, *ibid*. Rôle de l'accent latin dans la formation des —, *ibid*.
Nubie. Langues de la —, 93.
Oc (*Langue d'*). Dialectes de la —, 330.
Odji, 82.
Oïl (*Langue d'*). La —, au moyen âge, 327. Dialectes de la —, 328. — Limite actuelle de la — et de la langue d'oc, 329.
Omagua, 175.
Ombrien. Langue italique, 315.
Onéida, 175.
Omondago, 175.
Oráon, 114.
Osque. Langue italique, 315.
Ossète. Langue éranienne, 293.
Ostiaque. Samoyède —, 127. Parent du vogoul et du magyar, 134.
Otomi, 174.
Ottawa, 173.
Oude, 191.
Ouigour, 142.
Ouralo-altaïques (*Langues*). Se divisent en cinq grou-

pes, 125. Diversité des —, 126. Leurs caractères communs, *ibid.* L'incorporation dans les —, 139, 151. Harmonie vocalique des —, 152.
Padjadé, 80.
Paiamba. Dialecte australien. 95.
Pâli. Place du — parmi les idiomes prâkrits, 274.
Pani, 175.
Papous. Langue des —, 94.
Parsi. L'un des idiomes éraniens du moyen âge, 291.
Pâzend. Nom inexact donné au parsi, 291.
Pehlvi. Nom trop général donné au huzvârèche, 289.
Pépel, 80.
Permien. Langue ouralo-altaïque, 134.
Persan. La plus répandue des langues éraniennes modernes, 292. Sa grammaire, *ibid.*
Perse. Découverte du —, 284. Inscriptions cunéiformes —, *ibid.* Sa grammaire, 285.
Petit russe. Nom donné au ruthène, 370.
Phénicien. Appartient au groupe chananéen, 231. Est très rapproché de l'hébreu, 232. Le punique ou — d'Afrique, *ibid.*
Philologie. Distinction de la linguistique et de la —, 1. Ce que c'est que la — comparée, 2. Aide que la — peut attendre de la linguistique, 12.
Philologue. Distingué du linguiste, 7.
Phrygien. Langue indo-européenne de l'Asie Mineure, 398. Rapproché des langues éraniennes, *ibid.*
Physiologie. La — et la linguistique, 4.
Pikoumboul. Dialecte australien, 95.
Pluriels brisés. Les — du groupe arabe, 237, 241.
Polabe. Slave de l'Elbe, 382.
Polonais. Limites du —, 372. Sa grammaire, 373.
Polyglotte. Ne doit pas être confondu avec le linguiste, 14.
Polynésien. Groupe, 97.
Polysynthétisme. En quoi il diffère de l'incorporation proprement dite, 182.
Portugais. Distingué de l'espagnol, 334. Allié au galicien, *ibid.*
Poul. La langue —, 91. Distingue les êtres en deux catégories, 92.
Prâkrit. Relation du — au sanskrit, 274.
Primates. Tous les — précurseurs de l'homme n'acquirent pas la faculté du langage articulé, 37.
Provençal. Langue d'oc, 329. Sa période semi-analytique, 325.
Prussien (Vieux). Langue lettique disparue, 393.
Puetche, 176.
Punique. Phénicien d'Afrique, 233.
Quiché, 175.
Quichua, 176.
Race. Dans quelle condition la langue peut ne plus correspondre à la —, 422.
Racine. Définition de la —, 5. Dans les langues monosyllabiques la — constitue le mot 39. Dans quel but elle peut être modifiée par les langues à flexion, 206. Les — sémitiques sont réductibles, 212.
Râdjmâhal, 114.
Règne humain. La théorie du — est discréditée, 23.
Romaïque. Nom du grec moderne, 300.
Romanes (Langues). Nom donné aux langues novo-latines, 318. Hypothèse d'une langue romane, *ibid.*
Roumain. Sa place parmi les langues novo-latines, 335. Sa phonétique, 336. Son article, 337.

Roumanche. Ladin occidental, 333.
Rusniaque. Nom donné au ruthène, 370.
Russe. Limites du —, 369. Sa grammaire, 379.
Russe blanc. Dialecte — 369.
Ruthènes. Limites du —, 370. Comment il se différencie du russe, 371.
Saho. Dialecte éthiopien, 256.
Samoyède. Groupe — des langues ouralo-altaïques, 127.
Sanskrit. Premiers travaux sur le —, 271. Sa place dans la grammaire comparée des langues indo-européennes, *ibid*. Sa grammaire, 272. Son alphabet, 273. Sa littérature, 274.
Saxon (*Vieux*), 355.
Scandinave. Groupe —, 352.
Scythique. Prétendue langue —, 193. Quel sens faut-il attacher à ce mot ? 194.
Sémitique. La flexion — est autre que la flexion indo-européenne, 207. Le — et l'indo-européen sont irréductibles, 208. Les langues —, 210. Dénomination vicieuse, *ibid*. Du sémitisme en général, 211. Les racines — sont réductibles, 212. Le nom —, 213. Le verbe —, 215. Alphabet sémitique, 216. Classification des langues —, 217. Individualité des langues —, 242. Où a été parlée la langue — commune, 244. De quelle façon sont alliées les langues — et les langues khamitiques, 247.
Séna, 85.
Sénéca, 174.
Serbe. Voir **Serbo-croâte*, 377.
Serbo-croate Domaine du —, 377. Dialectes du —, *ibid*. Littérature —, 378. Importance du —, 379. Sa grammaire, 380.
Sérôlong, 86.
Sésouto, 86.
Sétchouana, 86.
Sétlapi, 86.
Siamois. Langue monosyllabique, 53.
Singhalais, 277.
Skipetar. Nom de l'albanais, 400.
Slave. Groupe—, 366. Limite des langues slaves au moyen âge, *ibid*. Alphabet —, 367. Langues slaves actuellement vivantes, 367. Classification des langues slaves, 386.
Slave ecclésiastique. Où fut parlé le —, 382.
Slovaque. Dialecte slave allié au tchèque, 374.
Slovène. Idiome sud-slave, 374.
Sokoto. Dialecte poul, 91.
Somâli. Dialecte éthiopien, 256.
Sondéen, 100.
Sonraï, 80.
Sorabe. Nom donné au slave de Lusace, 376.
Sorbe. Serbe de Lusace, 376.
Sousou, 80.
Suane, 192.
Suédois. Sa place dans le groupe scandinave, 353.
Sumérien. Ce que quelques auteurs entendent par ce mot, 197.
Suomi. Son importance dans le groupe finnois, 129. Sa grammaire, 130.
Syntaxe. La — naît avant la grammaire proprement dite, 42.
Syriaque. Sa place dans le groupe araméen, 218.
Syro-arabe. Synonyme du sémitique, 211.
Tagala, 98.
Tamachek. Dialecte libyen, 253.
Tamoul. Importance du — dans le groupe des langues dravidiennes, 112. Possède un alphabet particulier, 123.
Tatar. Nom du groupe turc, 142. — de Crimée, 143.
Tavghi. Dialecte samoyède, 127.

Tchèque. Limites du —, son orthographe, sa grammaire, 374.
Tchérémisse, 133.
Tchoude, 129.
Tcherkesse, 191.
Tchetchenze, 191.
Tchouktche américain, 189.
Tchouktche asiatique, 189.
Tchouvache, 143.
Téda, 82.
Téhuelche, 176.
Tékéza, 86.
Temps. Les deux — de la langue sémitique commune, 215. Les deux — de l'indo-européen commun, 262.
Téné, 80.
Été, 85.
Thouche, 191.
Tibétain. Est une langue monosyllabique, 55.
Tigré. Parent du ghez, 242.
Toma, 80.
Ton. Importance du — dans les langues monosyllabiques, 42. Différents — du chinois, 47. Différents — de l'annamite, 53. Différents — du siamois, 54. Différents — du birman, *ibid*.
Tongouse. Groupe — des langues ouralo-altaïques, 148.
Totonak. 175.
Touda, 114.
Toulou. Grand nombre des formes dérivées en —, 121.
Toumalé, 93.
Touranien. Mot vide de sens, 201. Théorie fallacieuse des langues touraniennes, *ibid*.
Tsaconien. Dialecte — du grec moderne, 300.
Tsiganes. Dialectes des —, 278.
Tudesque. Nom donné au vieux haut-allemand, 363.
Tupi, 175.
Turc. Goupe des langues ouralo-altaïques, 134. Grammaire du —, 136.
Turcoman, 134.
Tuskarora, 175.
Vepse, 129.
Verbe. Le — sémitique, 210. Le — indo-européen, 261.
Vinde. Nom donné au slave de Lusace, 372.
Vogoul, 134.
Vote, 129.
Votiaque, 134.
Wallon. Dialecte de la langue d'oïl, 328.
Wiratourroi. Dialecte australien, 95.
Wolaroi. Dialecte australien, 95.
Wolof, 75.
Yakout, 142.
Yorouba, 82.
Youkaghir, 189.
Yourak. Dialecte samoyède, 127.
Yuma, 175.
Zend. Découverte du —, 280. Le — est l'éranien oriental, 281. Sa grammaire, 282. Son âge, *ibid*.
Zoulou, 85.
Zuriène. Langue ouralo-altaïque, 134.

IMPRIMERIE DE CHOISY-LE-ROI

La Bibliothèque
Université d'Ottawa
Échéance

The Library
University of Ottawa
Date due

18 10 72

29 01 73

AUG 25 '81

AUG 10 '81

07 JAN. 1991

08 JAN. 1991

02 FEV. 1992

02 FEV. 1992

a39003 001561645b

CE P 0121
.H68 1911
C00 HOVELACQUE, LINGUISTIQ
ACC# 1381321

U D' / OF OTTAWA
COLL ROW MODULE SHELF BOX POS C
333 02 09 01 07 06 2